

APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaifanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1204
LENGTH: 352
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-1204

Query Match 10.7%; Score 69.6; DB 4; Length 352;
Best Local Similarity 56.6%; Pred. No. 1.6e-07;
Matches 129; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 23 ATGACACGCGGTATCCCGGAGAGACACCCGCGCCCGACGAGAGCCCGGTC 82
Db 123 ATGCACAGAGAAACATGAGGTGGCTGTGCTGGGGGACCCCGACACATCTTCCA 182
Qy 83 CACACAGCCCTCACACTGGGCGCCCGACCCCGCTCGAGACCACTGATCTGTCG 142
Db 183 AGGTCCACCGGTATCAATCACAAGCGAGACCTCCGCGCCGACATGCTGTGTC 242
Qy 143 GTGTTAGACCCCTTACCTGATCTGTGCTGCTGCGCTCCGCGGCTGCTACTCC 202
Db 243 CTGTTACACCCCTTCTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Qy 203 ATCAAGCCCGAGATCAGAGGTGTGTGATCACTGAGAGCGCCCG 250
Db 303 GTGAAGTCTAGGACAGAGAAATGTTGGCGACGTGACCGGCGCCG 350

RESULT 3
US-09-736-457-1204
Sequence 1204, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaifanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1204
LENGTH: 352
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-1204

Query Match 10.7%; Score 69.6; DB 4; Length 352;
Best Local Similarity 56.6%; Pred. No. 1.6e-07;

Matches 129; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
Qy 23 ATGACACGCGGTATCCCGGAGAGACACCCGCGCCCGACGAGAGCCCGGTC 82
Db 123 ATGCACAGAGAAACATGAGGTGGCTGTGCTGGGGGACCCCGACACATCTTCCA 182
Qy 83 CACACAGCCCTCACACTGGGCGCCCGACCCCGCTCGAGACCACTGATCTGTCG 142
Db 183 AGGTCCACCGGTATCAATCACAAGCGAGACCTCCGCGCCGACATGCTGTGTC 242
Qy 143 GTGTTAGACCCCTTACCTGATCTGTGCTGCTGCGCTCCGCGGCTGCTACTCC 202
Db 243 CTGTTACACCCCTTCTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Qy 203 ATCAAGCCCGAGATCAGAGGTGTGTGATCACTGAGAGCGCCCG 250
Db 303 GTGAAGTCTAGGACAGAGAAATGTTGGCGACGTGACCGGCGCCG 350

RESULT 4
US-09-702-705-998
Sequence 998, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaifanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 998
LENGTH: 207
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc Feature
LOCATION: (1) ... (207)
OTHER INFORMATION: n = A, T, C or G
US-09-702-705-998

Query Match 10.4%; Score 68; DB 4; Length 207;
Best Local Similarity 71.8%; Pred. No. 3.4e-07;
Matches 89; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 125 GACCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164
Db 83 GACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142
Qy 185 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 244
Db 143 ATGACATTCGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
Qy 245 GCCC 248
Db 203 GCCC 206

RESULT 5
US-09-736-457-998
Sequence 998, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong


```
RESULT 11
US-08-036-555B-163
; Sequence 163, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Mao Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; TITLE OF INVENTION: Preparation and Use
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Teal, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-036-555B-163

Query Match      6.6%; Score 43.2; DB 1; Length 745;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 246 CCGGCGGTTTGGCTCCAAAGCAAGTGTACAACATCTGCGCGGAGTGAAGCTGG 305
DB 29 CCGGCGGTTTGGCTCCAAAGCAAGTGTACAACATCTGCGCGGAGTGAAGCTGG 88
QY 306 TGGCGCACTGCTGCTCTGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 365
DB 89 TGGCGCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 148
QY 366 CCAAGGACTCTGCGCG 381
```

```
DB 149 CCGGCAACGAGCGCGC 164

RESULT 12
US-08-469-569-163
; Sequence 163, Application US/08469569
; Patent No. 5606032
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Mao Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; TITLE OF INVENTION: Preparation and Use
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,569
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Teal, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-469-569-163

Query Match      6.6%; Score 43.2; DB 1; Length 745;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 246 CCGGCGGTTTGGCTCCAAAGCAAGTGTACAACATCTGCGCGGAGTGAAGCTGG 305
DB 29 CCGGCGGTTTGGCTCCAAAGCAAGTGTACAACATCTGCGCGGAGTGAAGCTGG 88
QY 306 TGGCGCACTGCTGCTCTGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 365
DB 89 TGGCGCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 148
```

QY 366 CCAAGAGACTGCGGC 381
Db 149 CCGGCAACGAGCGGC 164

RESULT 13

US-08-249-322A-163
Sequence 163, Application US/08249322A
Patent No. 5716930
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
APPLICANT: Chen, Maio Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,322A
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION/DOCKET NUMBER: 34,266
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 163:
SEQUENCE CHARACTERISTICS:
LENGTH: 745
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-249-322A-163

Query Match 6.6%; Score 43.2; DB 1; Length 745;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 246 CCGGCGCTTTGGCTCCAAAGCAAGTGTACATCTGCGCGGATGTGACGCTGG 305
Db 29 CCGGCGCTTTGGCTCCAAAGCAAGTGTACATCTGCGCGGATGTGACGCTGG 88

QY 306 TGGCGCACTGTCTCTGAGGCTGTGAGTGAAGTGTGACCTGACCTGCGCGCTGG 365
Db 89 TGGCGCTGTGCTGCACACTACTCTGTCTGTGGAGACCGGCGCTTGGCGCGCGCGCG 148
QY 366 CCAAGAGACTGCGGC 381
Db 149 CCGGCAACGAGCGGC 164

RESULT 14

US-08-469-526A-163
Sequence 163, Application US/08469526A
Patent No. 5792849
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew
APPLICANT: Stroobant, Paul
APPLICANT: Minghetti, Luisa
APPLICANT: Waterfield, Michael
APPLICANT: Marchioni, Mark
APPLICANT: Chen, Maio Su
APPLICANT: Hiles, Ian
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
NUMBER OF SEQUENCES: 187
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,526A
FILING DATE: 06 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Beker-Brady, Kristina
REGISTRATION/DOCKET NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200A
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
INFORMATION FOR SEQ ID NO: 163:
SEQUENCE CHARACTERISTICS:
LENGTH: 745
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-526A-163

Query Match 6.6%; Score 43.2; DB 1; Length 745;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;


```

Db      520 ACCCAGCCGCC 531

RESULT 3
AV665659
LOCUS
DEFINITION
57 bp mRNA linear EST 28-NOV-2001
AV665659 Bos taurus brain fetus Bos taurus cDNA clone E1BR049H02
5', mRNA sequence.
ACCESSION
AV665659
VERSION
AV665659.1 GI:9924689
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
Takauega,A., Hirotsune,S., Itoh,R., Itohono,A., Suzuki,H., Aso,H.
1 (bases 1 to 557)
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
21570554
PUBMED
11713328
COMMENT
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazunugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers
source
1..557
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR049H02"
/issue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_1lb="Bos taurus brain fetus"
/notes="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT      89 a      214 c      158 g      91 t      5 others
ORIGIN
Query Match      52.1%; Score 339; DB 9; Length 557;
Best Local Similarity 83.9%; Pred. No. 3.8e-60;
Matches 407; Conservative 0; Mismatches 69; Indels 9; Gaps 2;

QY      9 CGCGCTGGAACCATGATGACGCGGATCCCGCGAGGACACCCGGGCCCCAGCGCCA 68
Db      12 CGAGCTGGAACCATGATGACGCGGATCCCGCGAGGACACCCGGGCCCCAGCGCCA 71
QY      69 GCAAGGCGG-----GTGCCACAGAGCCCTCACTGCGCGCCGCGACCCCGGCTC 122
Db      72 GCAAGGCGGAGCGGACAGCCGCGCTGAGCGCTGGGGGCTCCGCGACCCCGACCTC 131
QY      123 GAAACCACTGATCTGCTGCTGTTGAGGACCTCTTACCTGAATCTGTGCTCGGCT 182
Db      132 GAAACCACTGATCTGCTGCTGTTGAGGACCTCTTACCTGAATCTGTGCTCGGCT 191
QY      183 TCCTGGGCTGGCTACTCATCAAGGCGGAGATGAGAGTGTGAGCTGGAG 242
Db      192 TCCTGGGCTGGCTACTCATCAAGGCGGAGATGAGAGTGTGAGCTGGAG 251
QY      243 CGGCGGCGGCTTTTGCTCCAAAGCCAGGTCTACCAATCTGGCGCGAGTGGAGCG 302
Db      252 CAGCGCGGCTCTCGAGTCCAAAGCCAGGTCTACCAATCTGGCGCGAGTGGAGCG 311

```

```

QY      303 TGGTCCCGCCACTGCTGCTCTGAGGCTGTGTGTAAGTGTGCTCTGACACTGAGCCCGC 362
Db      312 TGGTCCCGCCCGGCTCTGCTCTGCTGTGTGTAAGTGTGCTCTGACACTGCTGTGGC 371
QY      363 TGGCAAGGACTGTGGCGGCTCTTCAAGACCAAGTTGATGACGGGATATGACTGAC 422
Db      372 TGGCAAGGACTGTGGCGGCTCTTCAAGACCAAGTTGATGACGGGATATGACTGAC 431
QY      423 AGGCTGGGCTCTGATCTGAGGCACTAGCCCGAGACACTGACCGGCTGCTGCCCTG 482
Db      432 AGGCGGGGCTCTGCTGCTG---TCTGACAGGAGACCCCAACCCGAGGCTTACCCAG 488
QY      483 GGGCC 487
Db      489 CCCCC 493

RESULT 4
LOCUS
DEFINITION
541 bp mRNA linear EST 04-OCT-2001
B1849231
ACCESSION
B1849231
VERSION
B1849231.1 GI:15961750
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 541)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casaas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karaymcheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
PUBMED
11282978
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 100 row: H column: 8
Seq primer: ATTTAGTGAACCTATAG.
Location/Qualifiers
source
1..541
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/issue_type="pooled"
/lab_host="DH10B"
/clone_1lb="MARC 2BOV"
/notes="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, thymus,
semiteudonous muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT      87 a      214 c      151 g      89 t
ORIGIN
Query Match      51.4%; Score 334.6; DB 12; Length 541;
Best Local Similarity 84.6%; Pred. No. 3e-59;
Matches 402; Conservative 0; Mismatches 64; Indels 9; Gaps 2;

```

QY 19 ACCATGACACGCGCTATCCCGGAGACACCCGCGCCCGCAGCAAGCGCG- 77
 DB 1 ACCATGACACGCGCTATCCCGGAGACACCCGCGCCCGCAGCAAGCGCGA 60
 QY 78 -----GTGCCACACAGCGCTTACATGCGCGCGCCCGCAGCGCGCTT 132
 DB 61 CGGACACGCGCCACACGCGCTGACGCTGGGGGCTCCCGACCCCGCAGCAAGCACTT 120
 QY 133 GATCTGCTGCTGCTTACAGCAACCTCTACCTGATCTGCTGCTGCTGCTGCTGCT 192
 DB 121 GATCTGCTGCTGCTTACAGCAACCTCTACCTGATCTGCTGCTGCTGCTGCTGCT 180
 QY 193 GGCTTACTCCATCAAGCGCGGAGATCAGAGGCTGCTGCTGCTGCTGCTGCTGCT 252
 DB 181 GGCTTACTCCATCAAGCGCGGAGATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 253 TTTTGGCTCCCAAGCGGCTACATCATCTGCGCGGATGAGAGCGCTGCTGCTGCTGCT 312
 DB 241 TCTGGCTCCCAAGCGGCTACATCATCTGCGCGGATGAGAGCGCTGCTGCTGCTGCTGCT 300
 QY 313 ACTGCTGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
 DB 301 GGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 373 CTGCTGCGCTTCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCT 432
 DB 361 CTGCTGCGCTTCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCT 420
 QY 433 CTGATCTGCGGCTACATGAGCGGAGACCTGACCGCGGCTGCTGCTGCTGCTGCTGCTGCT 487
 DB 421 CTGCTGCGCTGCTTCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCT 472

RESULT 5
 BM289220 434 bp mRNA linear EST 28-DEC-2001
 LOCUS 530774 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BM289220.1 GI:17998246
 VERSION EST.
 KEYWORDS Bos taurus (cow)
 SOURCE Bos taurus
 ORGANISM Bos taurus (cow)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G., Perle, G., Holt, I., Karaycheva, S., Liang, F., Quackenbush, J., and Keefe, J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)

TITLE
 JOURNAL MEDLINE PUBMED
 COMMENT
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithth@mail.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACCT
 BACKWARD: GTTCCAGTACGACG
 Plate: 144 row: M column: 3
 Seq primer: ATTGAGTACCTATG.
 Location/Qualifiers
 1..434

BASE COUNT 66 a 166 c 126 g 76 t
 ORIGIN
 Query Match 50.6%; Score 329.6; DB 12; Length 434;
 Best Local Similarity 88.2%; Pred. No. 3.1e-58;
 Matches 372; Conservative 0; Mismatches 44; Indels 6; Gaps 1;
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_11b="MARC 3BOV"
 /note="Vector: PCMV SPOT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."
 20 CCATGACACGCGCTATCCCGGAGACACCCGCGCCCGCAGCAAGCGCG-- 77
 DB 1 CCATGACACGCTGTAACCCCGGAGACCCCGGCGCCGACGCGCAGAGCGCGAC 60
 QY 78 -----GTGCCACACAGCGCTTACATGCGCGCGCCCGCAGCGCGCTT 133
 DB 61 GGCAAGCGCCACACGCGCTTACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 QY 134 ATCTGCTGCTGCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCT 193
 DB 121 ATCTGCTGCTGCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCTGCTGCT 180
 QY 194 GCTTACTCCATCAAGCGCGGAGATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253
 DB 181 GCTTACTCCATCAAGCGCGGAGATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 254 TTTGCTCCCAAGCGGCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCT 313
 DB 241 CTGCGCTCCAAAGCGGCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCTGCT 300
 QY 314 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373
 DB 301 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 374 TCTGCGCGCTTCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCTGCT 433
 DB 361 TCTGCGCGCTTCTTACAGCAACCTTATGACGCGGACTGATGACGAGCTGCTGCTGCTGCTGCTGCT 420
 QY 434 TG 435
 DB 421 TG 422

RESULT 6
 BQ207465/c 720 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ207465
 DEFINITION UT-R-DYI-cmo-e-14-0-UT-81 UT-R-DYI Rattus norvegicus cDNA clone
 ACCESSION BQ207465.1 GI:20423930
 VERSION EST.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 720)
 Bonaldo, M.F., Lennon, G., and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL MEDLINE PUBMED
 COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: benton-soares@uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 normalized cartilage library cDNA library Preparation: M.B. Soares
 Lab Clone distribution: clones will be available through Research
 Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA=yes

FEATURES

SOURCE

Location/Qualifiers

1. 720

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-DY1-cno-e-14-0-UI"

/tissue_type="Cartilage"

/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/clone_id="UI-R-DY1"

/note="Organ: Femur and Tibia; Vector: pT73D-Pac
 (Pharmacia) with a modified polylinker; Site 1: Not I;
 Site 2: Eco RI; UI-R-DY1 is a normalized cDNA library
 containing the following tissue(s): Rat Cartilage from
 Femur and Tibia. The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was ligated to an EcoR I adaptor, digested with Not
 I, and cloned directionally into pT73-Pac vector. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is CTATGAGACG. The Rat
 cartilage tissue was provided by Dr Jeff Stevens at the
 University of Iowa.
 TAG_LIB=UI-R-DY1
 TAG_TISSUE=cartilage
 TAG_SEQ=CTATGAGACG"

BASE COUNT 153 a 173 c 226 g 163 t 5 others

Query Match 45.4%; Score 295.4; DB 13; Length 720;
 Best Local Similarity 79.8%; Pred. No. 4e-51;
 Matches 360; Conservative 0; Mismatches 85; Indels 6; Gaps 1;

QY 1 GGAAGAGACGCGCTGGAACCATGACACGCGCTATCCCGCAGAGACACCGCGGCC 60
 DB 720 GGAAGACACGCGCTGGAACCATGACACCTTATCNCCTGAGAGACCCCGGCC 661
 QY 61 CAGCGCCAGCAAGCCCGT-----GCCACACAGCCCTTCACTGCGCGCCGCAACC 114
 DB 660 ATCATCCCGAAGGNTATGTCAGGCCACACGCGCCCTCTGTGAGACACCCGCGCC 601
 QY 115 CCGCGCTCGAGACCACTTGATCTGCTGGGTTCAGACCCCTTACCTGATCTGTGTTG 174
 DB 600 TAGCCACAGACCAATATGCTGTGCTGTTTCAAGCAGCTGATCTGATCTGTCTG 541
 QY 175 CTTGGCTTCTGCGCTGCGCTTCACTATCAGAGCCCGAGATCAGAGGTGTTGTGA 234
 DB 540 CTTGTTCTTCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
 QY 235 CTTGGAAGCGCCCGCGCTTGTGCTGCAAGCAAGTGTCTACACTCTGCGCGCGAT 294
 DB 480 CTTGGAAGCTGCAAGCAAGTGTGCTGCTGCAAGCAAGTGTCTGCTGCTGCAAT 421
 QY 295 GTGAGCGTGTGCGCGCACTGTCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCACT 354

DB 420 GTGAGACAGTGTCGCCCATTCATTCCTGCTGAGACTGTGTGATGATCGCGCTTGACCT 361
 QY 355 GCGCGCGCTGCGCAAGAGACTCTGCGCGCTTCTTCAGACCAAGTTGATGAGCGGACTA 414
 DB 360 GTCCCGCTGCGCAAGAGACTCTGCGCGCTTCTTCAGACCAAGTTGATGAGAGACTA 301

QY 415 TGACTGACAGCGTGGTCTGATCTGGGGCA 445
 DB 300 TAATGAGAGTCTGAGCGCTGCTGACCA 270

RESULT 7

AK003363

LOCUS

DEFINITION 739 bp mRNA linear HTC 05-DEC-2002

AK003363 Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
 enriched library, clone:11000306 product:hemoipoiesis related
 membrane protein 1, full insert sequence.

AK003363 GI:128333982

AK003363 HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishii, K., Kiyosawa, H., Kondo, S., Yamamato, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuenli, P., Lewis, S., Matsumoto, Y., Nishida, I., Pesole, G.,
 Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldirelli, P., Barsh, G., Blake, J., Botfield, D., Boyle, C.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Mazerall, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, T., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyooka, K., Wang, K. H., Weitz, C., Whitlaker, C., Wilming, L.,
 Wyshaw-Borja, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.,
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

PUBMED	21085660
REFERENCE	11217851
AUTHORS	5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCES	Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 739) Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiroake,T., Hori,F., Imotani,K., Ishii,Y., Itch,M., Iwata,M., Kasekawa,T., Kato,H., Kawai,J., Kojima,Y., Komio,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishik,K., Nomura,K., Numasaki,R., Ohno,W., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (10-JULY-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama Kanagawa 230-0045, Japan [E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216] Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAAGCGCCGCCCACTGACTTTTTCCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with full-length by cap-trapper sequence[5'-GAGAGAGAAGATCCAGACGTCAATTAAATTAATAAACCCCCCCC 3']. cDNA was cleaved with XhoI and NotI. Cloning sites, 5' end: SalI, 3' end: XhoI. Host: SOLR.
FEATURES	Location/Qualifiers 1..739 /organism="Mus musculus" /mol_type="mRNA" /strain="CS7BL/6J" /db_xref="PANTOM:DB:111000306" /db_xref="MGI:1906873" /db_xref="taxon:10090" /clone="111000306" /tissue_type="whole body" /clone_id="RIKEN full-length enriched mouse cDNA library /dev_stage="18-day embryo" 62..466 /note="haemopoiesis related membrane protein 1 [MGI] MGI:1934923, GB AJ009781, evidence: BLASTN, 99%, match=469) putative" /db_xref="MGI:1934923" 721..726 /note="putative" 739 /note="putative"
BASE COUNT	160 a 228 c 186 g 165 t
ORIGIN	
Query Match	43.9% Score 286; DB 11; Length 739;
Best Local Similarity	68.1%; Pred. No. 3.5e-49;
Matches	436; Conservative 0; Mismatches 185; Indels 19; Gaps 2
1	GGAAGAGACGGCGCTGGAAACCATGACACGCGCATATCCCGGAGAGACACCGGAGCCCC 60

[illegible]

source

1. .743
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30287396"
 /tissue_type="embryonic limb, maxilla and mandible"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH MGC 136"
 /note="Vector: pCMV-SPORT6.1.ccd; Site 1: EcoRV, Site 2: NotI; Normalized, full-length enriched library from pool of mouse embryonic limb, maxilla and mandible (day 17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw equivalents from respective days). Cloned directionally. oligo-dt primed (5'-GACTGACTTCAATCCGACGCGCCGCC(17)5-3'. Size selected for the 5kb fragments, average insert size 1.2 kb. Normalization to Cot 7.5. Tissue contributed by David Rowe, library constructed by Resgen, Invitrogen Corp. Note: this is a NIH MGC library."

BASE COUNT 167 a 226 c 176 g 162 t 12 others

ORIGIN

Query Match 43.7%; Score 284.4; DB 14; Length 743;
 Best Local Similarity 68.0%; Pred. No. 7.6e-49;
 Matches 435; Conservative 0; Mismatches 186; Indels 19; Gaps 2;

QY 1 GGAAGAGACGGCCCTGGAACCCATGACAGCGGCTATCCCGGAGACACCCGGGCC 60
 Db 16 GGAAGACACGGCTCTGGAACCCATGACATTCATCCCGGAGACACCCGGGCCCTCC 75
 QY 61 CAGCCACAGACGGCCGCT-----GCCACACAGCCCTCACTGGCGGCCCGGACCC 114
 Db 76 ATATATCCGAGAGGCTGATCTGACAGCCACACAGCCCTCTCAATGGAAACACTGGCCC 135
 QY 115 CCGGCTGAGACCTTGAATCTGTGTGATTCAGACACCCCTTACCTGATCTGTG 174
 Db 136 TACACACAGATATCATGCTGTGTCTTCAAGACAGATGATCTGATCTGTG 195
 QY 175 CCGGCTGCTTCTGGCGCTGCTCATCAAGCGCCGATCAGAGAGTGTGTG 234
 Db 196 CTTGATTCCTGCGCTGTCTCAAGCGCCGATCAGAGAGTGTGTG 255
 QY 235 CTTGAGAGCGCGCGCTTGTGCTCAAGCGAGTGTATACATCTGCGCGCGAT 294
 Db 256 CTTGAGAGCGCGCGCTTGTGCTCAAGCGAGTGTATACATCTGCGCGAT 315
 QY 295 GTGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
 Db 316 GTGAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 375
 QY 355 GCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 414
 Db 376 GTCAAGTTAGCAAGACTCTGCGCTTCTTCAAGCAAGTTGATGAGAGACTA 435
 QY 415 TGATGACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
 Db 436 TAACTAGAGTTTCCAGACTG-----TCTGTAACCGAGGAACATGCTCA 482
 QY 475 TGCCTGCGGCGCAATATGACTCTCCCGGAGCTGCGCTCTCTGCTGCGCTCAT 534
 Db 483 GGTCACTGCGCAACACAGAGCTCTCGGAGTTACAGCTTATAGACACCTGATCC 542
 QY 535 CCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
 Db 543 TGCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602
 QY 595 AGTCAGACCCCGGCTGCTTCAACCTTAACCGAGAGTTCC 634
 Db 603 ACTCAATCTCTCAACTTAATTAACCTTAAGGCTCC 642

LOCUS AA033499 485 bp mRNA linear EST 22-AUG-1996
 DEFINITION m144b03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 clone IMAGE:466349 5' similar to SW:INIB RAT P26376
 INTERFERON-INDUCIBLE PROTEIN. [1] ; mRNA sequence.
 ACCESSION AA033499
 VERSION AA033499.1 GI:1504910
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 485)
 REFERENCE
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished
 COMMENT Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:280165
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 404.
 Location/Qualifiers
 1. 485
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:466349"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /clone_1ib="Soares mouse embryo NbME13.5 14.5"
 /note="Vector: pT73D-Pac (pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCATCATGAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
 State Univ., from 2 1); double-stranded cDNA was ligated to
 Eco RI adaptors (pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M. Fatima Bernaldo."

BASE COUNT 108 a 142 c 133 g 102 t

ORIGIN

Query Match 43.5%; Score 283.2; DB 9; Length 485;
 Best Local Similarity 80.5%; Pred. No. 1.2e-48;
 Matches 346; Conservative 0; Mismatches 78; Indels 6; Gaps 1;

QY 1 GGAAGAGACGGCGCTGGAACCCATGACAGCGGCTATCCCGGAGACACCCGGGCC 60
 Db 51 GGAAGACACGGCGCTGGAACCCATGACATTCATCCCGGAGACACCCGGGCCCTCC 110
 QY 61 CAGCCACAGACGGCCGCT-----GCCACACAGCCCTCACTGGCGGCCCGGACCC 114
 Db 111 ATATATCCGAGAGGCTGATCTGACAGCCACAGCCCTCTTCAATGGAAACCTGGGCC 170
 QY 115 CCGGCTGAGACCTTGAATCTGTGTGATTCAGACACCCCTTACCTGATCTGTG 174
 Db 171 TACACACAGATATCATGCTGTGTCTTCAAGACAGATGATCACTGAATCTGTGCTG 230

QY	175	CCTGGGCTTCCTGGCGCTGGGCTACTCATCAAGGCCGAAGTACGAAGTGTGGGTA	234
Db	231	CCTTGGATCTCTGGGCGCTGGTCCACTCTGTCAAGGCCCGAGACCAAGAAATGGCTGGGAA	290
QY	235	CCTGGAAAGCGGCCCGGCGGTTTTGGCTCCAAAGCCAAAGTGTACAAACATCTGGCCGCGAT	294
Db	291	CTTGGAGGCTGCCAAGCGACGATATGGCTCCAAAGCCAAAGTGTACAAACATCTGGCTGCAGAT	350
QY	295	GTGACGCTGTGGACGCGACACTGCTGCTCTGGAGGCTGGTGGTGACTGGTGGCCCTGCACCT	354
Db	351	GTGACATTTGGTGGCCCATTTGCTGCTCTTGGGACTGGTGGTGACTGGCGGCTTTGGACCT	410
QY	355	GGCCCGGCTGGCCAAAGACTCTGCGCGCTTTCTTGACCAACAATTGATGAGCGCGACTA	414
Db	411	GTCCAAAGTTAACCAAAAGACTCTGGCGGCTTCTTGACGACCAAGTTGATGAGAGGACTA	470
QY	415	TGACTGACAG 424	
Db	471	TAACTAAGAG 480	

RESULT 10	BY703620	LOCUS	DEFINITION	LOCUS	DEFINITION
	BY703620	683 bp	mRNA	linear	EST 16-DEC-2000
	BY703620	RIKEN full-length enriched, 18-day embryo	whole body	Mus	
		musculus cDNA clone 111000J06 5', mRNA sequence.			
		111000J06			

ACCESSION	BI 703620
VERSION	BY703620.1
KEYWORDS	GI:27114731
	NOT

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE AUTHORS

TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)

MEDLINE	22354683
PUBMED	12466851

COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-ree@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda,
S., Hashizume, W., Hayashida, K., Hirose, T., Hoti, F., Imocani, K.,
Ishii, Y., Itoh, M., Kageana, I., Kawai, O., Kojima, S., Komio,
H., Koyu, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,
Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y.: Direct
Substitution
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

SOURCE

Location/Qualifiers
1. .683

BASE COUNT
ORIGIN

150 a 205 c 174 g 154 t

Query March Similarity	43.5%	Score 283;	DB 14;	Length 683;
Best Local Similarity	79.4%	Pred. No. 1.4e-48;		
Matches 350;	Conservative	0;	Mismatches 85;	Indels 6;
				Gaps 1;

QY 1 GGAGAGACGCGGCTGGAAACCATGAGACACGCGGTATCCCCGGAGGACACCCGGGGCCCC 60
Db 40 GGAGAGACACGCGCTGGAAACCATGAGACATTATATATCCCCGGAGGACCCCCCGGGCTCC 99

61 CACGCCCAAGCAGCGCGGT-----GCCACACAGCCTTCACACTGCGCGGCCCCCGACCC 114
OY
100 ATCATCCCGCAGAGGCTGTGCTGCAGGCCACACAGCCCTTCATGGGAAACACCTTGCCC 159
bb

QY 115 CCGCGCTGAGACCACTTGATCTGGTCGGTGTTCAGCACCCCTTACTCGAATCTGTGTTG 174
 DB 160 TACACCAAGAGTTCATGCTCTGGTCTGTCTTCAGCAGCATGTACCTGGAATCTGTGCTG 219

175 CCGCGCTTCCTGCGCGCTGCGCTACTCATCAAGGCCGAGATCAGAAAGTGTTGTGA 234
 220 CCTTGATTTCTGGGCGCTGATCCACTGTGTCAAGGCCGAGA CTAGAAGATGGCTGGGA 279

235 CCTGGAAGCGGCCCGCGCTTTTGAGCTCCAAAGCCAAAGTGTACACACATCTGCGCGCAT 294
 280 CTTTGGAGGTTTGTATAGGTAAGTAAAGTCTCTTAAAGCTTAAGTGTATACATCTTGGCTGCAT 339

[illegible]

355 GGCCCGCTGGCAAGACTCTGCGCGCTTCTTCAGCACCAGTTTGATGACCGGACTTA 414

QY	415	TC	ACTG	CA	GG	CTG	GG	TT	CCTG	435
Db	460	TA	ACTA	GA	GT	CC	GA	CC	CTG	480

RESULT 11
BQ192887/c

LOCUS	B01936887	725 bp	EST 30-APR-2002
DEFINITION	UI-R-DRI-clk-m-07-0-UI s1	UI-R-DRI	Rattus norvegicus cDNA clone
ACCESSION	UI-R-DRI-clk-m-07-0-UI 3'	mRNA sequence.	
UNIQUE	B01936887		
CLONE	B01936887		

SOURCE ORGANISM	Rattus norvegicus (Norway rat) Rattus norvegicus

REFERENCE	AUTHORS	TITLE
1 (pages 1 to 725)	Bonaldo, M.F., Lennon, G. and Soares, M.B.	Normalization and subtraction: two approaches to facilitate gene

COMMENT **Contact: Soares, MB**

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized osteoblast library cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.researchgen.com)
Seq primer: M13 Forward
POLYA=yes.

FEATURES	Location/Qualifiers
source	1. .725

```

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DR1-clk-m-07-0-UI"
/seq_range="adult"
/lab_host="DH10B (Life Technologies)"
/clone_idb="UI-R-DR1"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DR1
library is a normalized Rat Osteoblast library (NRRO)
constructed in pT73 vector according to the procedure
described by Bonaldi, Lennon & Soares (Normalization and
Subcloning: Two Approaches to Facilitate Gene Discovery
Genome Research 6: 791-806, 1996). The oligonucleotide
used to prime first strand synthesis contained the
sequence tag AACATTCACA between the Not I cloning site and
d18 stretch. The Rat Osteoblast tissue was provided by
Lian & Stein of the University of Massachusetts Medical
School.

```

BASE COUNT	155 a	175 c	228 g	165 t	2 others
ORIGIN					

Query Match	43.0%	Score 280;	DB 13;	Length 725;
Best Local Similarity	79.4%;	Pred. No. 6.1e-48;		

Matches 358; Conservative 0; Mismatches 86; Indels 7; Gaps 2;

QY	1	GGAAAGACGGGGCTGGAAACCCATGAGACAGCGCGTATCCCGGAGAGACACC	GGGGCCCC	60
Db	719	GGAAAGAGGGGGCGCTGGAAACCATGACATTCATATCCCCGTGAGACAC	CCGGCCCC	660
QY	61	CACGCCACGAAAGCCGGT-----GCCCAACAGCCCTCACA	CTGGCGCCCCGACC	114
Db	659	ATCATCCCGCAAGGCTGATGATGTGACAGCCCA	CAGGCCCTCTCTGTGG--GACACCCGGCCC	601
QY	115	CCCGGCTGGAACCACTTGATCTGGTCGGTTCGGTTCAGACACCTCTACCTGAATCTGTGTTG	174	
Db	600	TACGCAAGAACCAATATGCTCTGTCTGTCTTTCAGACAGCTGATCTGAATCTGTGCTG	541	
QY	175	CCTCGGCTTCTCGGGCGCTGCGCTACTCATCAAGCCCGAGATCAGAAGTGTGTGTGA	234	
Db	540	CCTTGTGTTTCTCGGGCGCTGGTGCATCTGTGTCAAGCCCGAGACAGAAAGTGA	CCGGGA	481
QY	235	CCTGGAAGCGCCCGCGCGTTTGTGCTCAAAAGCCAAAGTGTCTAACATCTCTGGCCGCA	T	294
Db	480	CCTGGAAGCTCAAAAGCAAGTATGTGTCCAAAGCCAAAGTGTCTCAACATCTGGCTCAAT		421
QY	295	GTGACGCTGTGGTCCGCGACATGCTCTCTGTGGGGGTGGTGTGATGATGAGCCCTGCACCT		354
Db	420	GTGACATCTGGTGCCCCCATTTGCTCTCTGGGACTGGTGTGATCTGGGCGCTTGCACCT		361
QY	355	GGCCCGCTGGCCCAAGGACTCTGCCGCTTCTTCAGACCAAGTTTGATGACGCGACTA		414
Db	360	GTCCCGGCTAACCAAAAGACTGTGGCGGCTTTCTTCAGACCAAGTTTGATGAGGAGACTA		301
QY	415	TGACTGACAGGCTGGTCTGATCTGGGGCA		445
Db	300	TAACTGAGAGTTCTGAGGCTGTCTCTGAACCA		270

RESULT 12

LOCUS	738 bp	mrna	linear	EST	04-MAR-2003
DEFINITION	CB321092	AGENCOURT 12238330	NIH MGC_136	Mus musculus	CDNA clone
IMAGE:	30258864	5',	mrna	sequence.	

ACCESSION	CB321093
VERSION	CB321093.1
KEYWORDS	GI:28845328
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus 1 (Bases 1 to 738)
TITLE	NIH-MGC http://mgc.nhi.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished
	Contact: Robert Strausberg, Ph.D.

FEATURES

BOURCE

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3028864"
/tissue_type="embryonic limb, maxilla and mandible"
/lab_note="DH10 (phage-resistant)"
/clone_lib="NH MCC 136"
note="Vector: pCMV-SPORT6.1.ccd; Site_1: EcoRV; Site_2:
NotI; Normalized, full-length enriched library from pool

```


FEATURES	source
495 GAGTCCCGGAGCCCGGCGCCCTCTCTGTGGGGCGCTCCATCCCTGCCCATCTGATCT 554	
468 AGCTCTGGAGATTACAGCTTACTATAGCACTTCATCCCTGCGCCAGCGTGAACCTGAGA 527	
555 GGGGCGCTCCAGCCCAACATGGGACACCTTAAGGTGAACCAATCCAGATCCCGGGGTCTTC 614	
528 GCGGCGCCTCTTGTTCACGATGTGTATTGATGTGCGCCCACTGAGATCCCTCAGACTTA 587	
615 ACCCTAACCCGAGAGTTCCC 634	
588 ACTATATAACCTAGAGGCTCC 607	
LOCUS	
AA000766	
DEFINITION	443 bp mRNA linear EST 18-JUL-1996
ACCESSION	U000766
KEYWORDS	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 443)
AUTHORS	Matra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Macerston, R.
TITLE	The WashU-HMNI Mouse EST Project
JOURNAL	Unpublished
COMMENT	Contact: Maxra W/Mouse EST Project WashU-HMNI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:260576 Seq primer: ETPRimer High quality sequence stop: 354. Location/Qualifiers 1..443 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:426024" /sex="unknown" /tissue_type="embryo" /dev_stage="13.5-14.5dpc total fetus" /lab_host="DHI0B" /clone_id="Soares mouse embryo NDME13.5 14.5" /note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15', TGTTACCAATCTGAAGTGGGAGCGCGGAGAAATTTTATTTTTTTTTTTTTTTT T 3'1', on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2 1]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldi. "
BASE COUNT	96 a 135 c 115 g 97 t
ORIGIN	

Query Match	40.2%	Score 261.6;	DB 9;	Length 443;
Best Local Similarity	78.5%	Pred. No. 3,4e-44;		
Matches 328;	Conservative 0;	Mismatches 86;	Indels 6;	Gaps 1.
QY	24	TGACACGGCGTATCCCGCGAGGACACCCGGGCCCCCGACCCGACGACGAGCCGGT----	79	
Db	3	TGGACACTTCATATCCCGGTATGATACCCCGGGCTCCATCATCCCGCAGGCTGATGCTG	62	
QY	80	--GCCACACAGACCCCTGCACATGGGGGGCCCGACACCCCGGCGCTGAGAACCATGTGACT	137	
Db	63	CAGCCCAACAGCCCTTCTCCATGGGACACCTGGCCCTTACACACAGATCATCATGCTCT	122	
QY	138	GCTGGGTTCAGACACCCCTACCTGATCTGTGTGTGCTCGGCTTCTCGGCGTCGACT	197	
Db	123	GGTCTGTCTTCAGACACGATGACTGATCTGTGTGTGCTTGGATTCCTGGGGCTGATCC	182	
QY	198	ACTCCATTAAGGCCCGGAGATCATGAAGGTGTGTGTGACTGTGAAGCGGCGCGCTTTTG	257	
Db	183	ACTGTCTAAGGCCCGGAGACCGAAGATGGCTGGGAATTGGAAGCTGTCAAGGACGATATG	242	
QY	258	GCTCCAAAGCAATGTCTACAACATCTCTGGCGCGAGATGTGACGCTGGTGGCCGACATGC	317	
Db	243	GCTCCAAAGCAATGTCTACAACATCTCTGGCTGCAATGTGACATGTGCTGCCCATTTGC	302	
QY	318	TGCTTCCTGGGGCTGGTGTGACTGTGTGCTCTGCACCTTGGCCCGGCTGGCCAAGACTGTG	377	
Db	303	TGCTTCCTGGGGCTGGTGTGACTGTGTGCTCTGCACCTGTCCAAATTGACCAAGACTGTG	362	
QY	378	CCGCGCTTTCAGGACCAAGTTTATATGACGGGACTATATGACAGGCTGGTCTTG	435	
Db	363	CGGCTTTCAGGACCAAGTTTATATGAGGAGACTATATGAAAGTTTCCGAGCTTG	420	

```

RESULT 15
LOCUS      W75342
DEFINITION W75342      462 bp      mRNA      linear      EST 20-JUN-1996
              me49ph03.r1 Soares mouse embryo NDME13.5 14.5 Mus musculus cDNA
              clone IMAGE390869 5' similar to SW:INIB RAT P26376
              INTERFERON-INDUCIBLE PROTEIN. [1] ;, mRNA sequence.
ACCESSION  W75342
VERSION    W75342.1
KEYWORDS   W75342.1 GI:1385557
SOURCE     EST.
           Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 462)
            Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellhammer,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Thaisang,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE      The WashU-HHMI Mouse EST Project
JOURNAL    Unpublished
COMMENT    Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine#
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@wustl.edu
            This clone is available royalty-free through LIAAT ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:242701
FEATURES   Seq primer: mob.BEGA+ET
            High quality sequence stop: 343.
            Location/Qualifiers
                1..462
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
FEATURES   source

```

```
/clone="IMAGE:390869"  
/sex="unknown"  
/tissue_type="embryo"  
/dev_stage="13.5-14.5dpc total fetus"  
/lab_host="DH10B"  
/clone_lib="Soares mouse embryo NDM3.5 14.5"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCGGAATTTTTTTTTTTTTTTT  
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2  
14.5dpc embryos (total RNA provided by Minoru Ko, Wayne  
State Univ., from 2 ); double-stranded cDNA was ligated to  
Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
pT73 vector. Library went through one round of  
normalization, and was constructed by Bento Soares and  
M.Patrina Bonaldo."
```

```
BASE COUNT      101 a      141 c      123 g      97 t  
ORIGIN
```

```
Query Match      40.2%; Score 261.6; DB 14; Length 462;  
Best Local Similarity 79.2%; Pred. No. 3.4e-44;  
Matches 350; Conservative 0; Mismatches 84; Indels 8; Gaps 3;
```

```
QY      1 GGAAGAGCGCGCTGGAACCATGACACGCGGTATCCCGCGAGAGACCCGCGCCC 60  
Db      5 GGAAGACACGCGCGCTGGAACCATGACATTCATATCCCGTGA-GACCCCGCGCTCC 63  
QY      61 CAGGCCAGACAGCGCGCT-----GCCACACAGCGCTCACTGCGCGCCCGCACCC 114  
Db      64 ATCATCCCGAGCGCTATGCTGACAGCCACAGCCCTCTCCATGGGAAACCTGCGCC 123  
QY      115 CCCGCTCGAGACCATGATCTGATGATGATGATGATGATGATGATGATGATG 174  
Db      124 TAGACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 183  
QY      175 CTTGCGCTTCTGCGCTGCGCTTACTCATCAGGCCCGAGATCAGAAAGTGTGTA 234  
Db      184 CTTGGAATCTCTGCGCTGCTGCTCACTGTCANAGGCCGAGACGAGAAATGCTGGAA 243  
QY      235 CTTGGAAGCGCGCGCGCTTGTGCTCAAA-GCCAGTGTACATCATCTGCGCGCA 293  
Db      244 CTTGGAAGCTGCAAGGAGATGATGATGATGATGATGATGATGATGATGATG 303  
QY      294 TGTGAGCGTGTGTCGCGCACTGCTGCTCTGCGGCTGTGTGATGTGCTGCACC 353  
Db      304 TGTGAGCATGTGTGTCGCGCACTGCTGCTCTGCGGCTGTGTGATGTGCTGCACC 363  
QY      354 TGGCCCGGCTGCGCAAGACTCTGCGCGCTTCTTCAAGACCAAGTTGATGACGGGACT 413  
Db      364 TGTCCAAGTTAGCCAAAGACTCTGCGCTTCTTCAAGACCAAGTTGATGAGGAGACT 423  
QY      414 ATGACTGACAGGCTGGGTCTTG 435  
Db      424 ATTAATAAGATTCCGAGCTTG 445
```

```
Search completed: January 5, 2004, 18:04:06  
Job time : 1977 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 16:30:42 ; Search time 2818 Seconds

(without alignments)
9450,734 Million cell updates/sec

Title: US-09-914-815A-10

Perfect score: 651

Sequence: 1 ggaagagacggcgctggaac.....ccggggccctaactctgcc 651

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_ey:*
13: gb_un:*
14: gb_vt:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vt:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_ey:*
39: em_hngo_hum:*
40: em_hngo_mus:*
41: em_hngo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	446.6	68.6	95038	9 AF015416	AF015416 Homo sapi
C 2	446.6	68.6	143835	2 AF015416	AF015416 Homo sapi
C 3	446.6	68.6	156819	2 AP006283	AP006283 Homo sapi
C 4	446.6	68.6	175416	2 AP006286	AP006286 Homo sapi
C 5	446.6	68.6	187160	2 AC138230	AC138230 Homo sapi
C 6	284.6	43.7	499	10 MMU9781	MMU9781 Mus muscu
C 7	173	26.6	238533	2 AC132968	AC132968 Rattus no
C 8	167.6	25.7	200734	2 AC109272	AC109272 Mus muscu
C 9	167.6	25.7	251206	2 AC107815	AC107815 Mus muscu
C 10	95.4	14.7	136098	2 AC006970	AC006970 Homo sapi
C 11	93.8	14.4	51891	2 AC074103	AC074103 Homo sapi
C 12	93.4	14.3	678	9 BC009696	BC009696 Homo sapi
C 13	93.4	14.3	905	6 AX337050	AX337050 Sequence
C 14	93.4	14.3	905	9 HS18D	HS18D Human 1-8D
C 15	93.2	14.3	402	9 BT006892	BT006892 Homo sapi
C 16	93.2	14.3	402	12 BT007876	BT007876 Synthetic
C 17	93.2	14.3	637	9 BC006794	BC006794 Homo sapi
C 18	93.2	14.3	642	6 AR225472	AR225472 Sequence
C 19	93.2	14.3	642	6 AX321588	AX321588 Sequence
C 20	93.2	14.3	645	9 BC022439	BC022439 Homo sapi
C 21	93.2	14.3	695	6 AX281855	AX281855 Sequence
C 22	93.2	14.3	749	9 BC008417	BC008417 Homo sapi
C 23	93.2	14.3	808	9 HS18U	HS18U Human 1-8U
C 24	91.6	14.1	37918	9 AP005232	AP005232 Homo sapi
C 25	91.6	14.1	222765	2 AC144988	AC144988 Gorilla g
C 26	91.2	14.0	378	9 BT007173	BT007173 Homo sapi
C 27	91.2	14.0	378	12 BT007946	BT007946 Synthetic
C 28	91.2	14.0	647	9 HSRNA927	HSRNA927 Homo sapi
C 29	91.2	14.0	683	9 BC000897	BC000897 Homo sapi
C 30	91.2	14.0	842	6 AX397528	AX397528 Sequence
C 31	91.2	14.0	851	6 AX014337	AX014337 Sequence
C 32	91.2	14.0	853	6 AX302539	AX302539 Sequence
C 33	91.2	14.0	853	6 AX409461	AX409461 Sequence
C 34	91.2	14.0	853	6 HOM927A	HOM927A Human inter
C 35	91.2	14.0	1095	6 AX281854	AX281854 Sequence
C 36	91.2	14.0	245	6 AX340862	AX340862 Sequence
C 37	90.8	13.9	579	4 AF272041	AF272041 Bos tauru
C 38	90.8	13.9	107717	9 AC137055	AC137055 Homo sapi
C 39	90.8	13.9	110000	2 AC004085_1	AC004085_1 Continu
C 40	90.8	13.9	110000	2 AC004085_2	AC004085_2 Continu
C 41	90.2	13.9	122351	9 AC023157	AC023157 Homo sapi
C 42	89.6	13.8	905	10 BC027285	BC027285 Mus muscu
C 43	89.6	13.8	84979	9 AC116049	AC116049 Homo sapi
C 44	89.6	13.8	176104	2 AC016193	AC016193 Homo sapi
C 45	89.6	13.8	189572	2 AC084812	AC084812 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AF015416/c 95038 bp DNA linear PRI 20-AUG-1997
DEFINITION Homo sapiens chromosome 11 from 11p15.5 region, complete sequence.
ACCESSION AF015416
VERSION AF015416.1 GI:2335202
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 95038)
Evans, G.A., Bradbury, P., Brignac, S., Bumeister, R., Davie, J.,
Davies, C.J., Davis, C., English, C., Fondon, T., Franklin, T.L.,
Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J.,

TITLE
JOURNAL
REFERENCE
AUTHORS
Hinson, S., Megarity, C., Narayanaswamy, U., Newton, J., O'Brien, K.,
Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J.,
Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T.,
Wilson, R. and Burbee, D.,
95 Kb from Four Overlapping Human Chromosome 11p15.5 Cosmids
Unpublished
2 (bases 1 to 95038)
Evans, G.A., Bradbury, P., Brignac, S., Bumeister, R., Davie, J.,
Davies, C.J., Davis, C., English, C., Fondon, T., Franklin, T.L.,
Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J.,
Hinson, S., Megarity, C., Narayanaswamy, U., Newton, J., O'Brien, K.,
Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J.,
Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T.,
Wilson, R. and Burbee, D.,
Direct Submission
TITLE
JOURNAL
Submitted (18-JUL-1997) Genome Science and Technology Center,
University of Texas Southwestern Medical Center at Dallas, 5323
Harry Hines Blvd, Dallas, TX 75235-8591, USA
FEATURES
SOURCE
1. .95038
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p15.5"
711. .975
repeat_region /rpt_family="Alu"
2728. .3022
repeat_region /rpt_family="Alu"
3363. .3719
repeat_region /rpt_family="THB1"
5634. .5909
repeat_region /rpt_family="Alu"
complement(6720. .6959)
/rpt_family="Alu"
10560. .10839
repeat_region /rpt_family="Alu"
complement(11832. .12455)
/rpt_family="Alu"
12821. .13405
repeat_region /rpt_family="L1"
13718. .14314
repeat_region /rpt_family="Alu"
14617. .14779
repeat_region /rpt_family="Alu"
15396. .15695
repeat_region /rpt_family="Alu"
15722. .15994
repeat_region /rpt_family="Alu"
complement(16696. .16997)
/rpt_family="Alu"
complement(18996. .19286)
/rpt_family="Alu"
complement(19458. .19747)
/rpt_family="Alu"
complement(19777. .20045)
/rpt_family="Alu"
complement(22885. .23002)
/rpt_family="MIR"
23109. .23261
repeat_region /rpt_family="Alu"
23292. .23481
repeat_region /rpt_family="MER46"
complement(23877. .23944)
/rpt_family="MIR"
complement(24695. .24986)
/rpt_family="Alu"
complement(25429. .26012)
/rpt_family="Alu"
complement(27380. .27676)
/rpt_family="Alu"
complement(29422. .29486)
/rpt_family="MER42"

repeat_region complement(29514. .29802)
/rpt_family="Alu"
30797. .31075
repeat_region /rpt_family="Alu"
31688. .31792
repeat_region /rpt_family="Alu"
32125. .32412
repeat_region /rpt_family="Alu"
complement(38871. .39186)
/rpt_family="Alu"
42952. .43094
repeat_region /rpt_family="MERS"
43106. .43388
repeat_region /rpt_family="Alu"
43754. .44216
repeat_region /rpt_family="MT2B2"
complement(44815. .45136)
/rpt_family="MER42"
45145. .45484
repeat_region /rpt_family="Alu"
45740. .45995
repeat_region /rpt_family="Alu"
46375. .46646
repeat_region /rpt_family="Alu"
47145. .47738
repeat_region /rpt_family="Alu"
48353. .48626
repeat_region /rpt_family="Alu"
49567. .49842
repeat_region /rpt_family="Alu"
49940. .50211
repeat_region /rpt_family="Alu"
complement(50615. .50907)
/rpt_family="Alu"
complement(53001. .53290)
/rpt_family="Alu"
complement(54161. .54512)
/rpt_family="L1"
54533. .54823
repeat_region /rpt_family="Alu"
57899. .58113
repeat_region /rpt_family="Alu"
complement(58795. .59044)
/rpt_family="Alu"
complement(59414. .60023)
/rpt_family="L1"
complement(60177. .60311)
/rpt_family="Alu"
complement(60770. .61108)
/rpt_family="L1"
complement(61207. .61557)
/rpt_family="THE1"
61670. .61942
repeat_region /rpt_family="Alu"
61944. .62436
repeat_region /rpt_family="MER41"
complement(67807. .68105)
/rpt_family="Alu"
70961. .71049
repeat_region /rpt_family="Alu"
71224. .71523
repeat_region /rpt_family="Alu"
complement(85033. .85397)
/rpt_family="MER4"
complement(85414. .85620)
/rpt_family="Alu"
complement(85701. .85804)
/rpt_family="MER4"
complement(85895. .86028)
/rpt_family="MER4"
86180. .86477
repeat_region /rpt_family="Alu"
86497. .86780
repeat_region

```

repeat_region      /rpt_family="Alu"
                    87125..87256
repeat_region      /rpt_family="Alu"
                    87590..87893
repeat_region      /rpt_family="Alu"
                    89103..89401
repeat_region      /rpt_family="Alu"
                    89726..89837
repeat_region      /rpt_family="Alu"
                    89888..90390
repeat_region      /rpt_family="MER4"
                    complement(91891)
repeat_region      /rpt_family="Alu"
                    complement(93951..94017)
repeat_region      /rpt_family="L1R12"
                    complement(93951..94017)
BASE COUNT      23552 a 24227 c 24425 g 22834 t
ORIGIN
Query Match      68.6%; Score 446.6; DB 9; Length 95038;
Best Local Similarity 99.1%; Pred. No. 2.5e-76;
Matches 445; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY      199 CTCCTCAAGGCGCCGAGATCAGAGTGTGTGACCTGGAAGCGGCGGCTTTGG 258
Db      83096 CTCCTCGAGGCGCGAGATCGAGAGTGTGTGACCTGGAAGCGGCGGCTTTGG 83037
QY      259 CTCCTCAAGGCGCCGAGATCAGAGTGTGTGACCTGGAAGCGGCGGCTTTGG 318
Db      83036 CTCCTCAAGGCGCCGAGATCAGAGTGTGTGACCTGGAAGCGGCGGCTTTGG 82977
QY      319 GCTCTGTGGGCGTGTGTGATGCTGTGCGCTGACCTGCGCGGCTGGCCAGAGCTTGC 378
Db      82976 GCTCTGTGGGCGTGTGTGATGCTGTGCGCTGACCTGCGCGGCTGGCCAGAGCTTGC 82917
QY      379 CGCCTTCTTACGACCAAGTTTGTGACGCGGAGTATGACTGACGAGCTGGGCTTGC 438
Db      82916 CGCCTTCTTACGACCAAGTTTGTGACGCGGAGTATGACTGACGAGCTGGGCTTGC 82857
QY      439 TGGGGGCACTAGGCGCCGAGACCTGACCCGAGCTGTGCGGCGGCGGCAATAGTACT 498
Db      82856 TGGGGGCACTAGGCGCCGAGACCTGACCCGAGCTGTGCGGCGGCGGCAATAGTACT 82797
QY      499 CCGCGGAGCGTGGCGGCTTCTGTGTGGGCGCTCCATCCCTGCGGCTTGTATCTTGGGG 558
Db      82796 CCGCGGAGCGTGGCGGCTTCTGTGTGGGCGCTCCATCCCTGCGGCTTGTATCTTGGGG 82737
QY      559 CCGCTCAGCCCAACATGAGGAGCTTAAGGCTGAACAGTCAAGCCCGGGGCTTTCAACC 618
Db      82736 CCGCTCAGCCCAACATGAGGAGCTTAAGGCTGAACAGTCAAGCCCGGGGCTTTCAACC 82677
QY      619 TAACCCGAGAGTTCCCGGGCGCTTAAGTCTTGGCC 651
Db      82676 TAACCCGAGAGTTCCCGGGCGCTTAAGTCTTGGCC 82644

RESULT 2
AC136475/c      143835 bp      DNA      linear      PRI 30-MAY-2003
LOCUS      AC136475
DEFINITION      Homo sapiens chromosome 11, clone RP11-326C3, complete sequence.
ACCESSION      AC136475
VERSION      AC136475.6
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 143835)
AUTHORS      Birren, B., Nusbaum, C. and Lander, E.
TITLE      Homo sapiens chromosome 11, clone RP11-326C3
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 143835)
AUTHORS      Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B.,

```

```

TITLE      JOURNAL
COMMENT
REFERENCE
AUTHORS
TITLE      JOURNAL
COMMENT
Submitted (02-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 143835)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Boguslavsky, L., Bouhgalter, B., Barna, N., Bastien, V., Bloom, T.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzderald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, T., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menus, L., Mihova, T., Mlewa, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachpka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 143835)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Boguslavsky, L., Bouhgalter, B., Barna, N., Bastien, V., Bloom, T.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzderald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, T., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menus, L., Mihova, T., Mlewa, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachpka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (30-MAY-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 30, 2003 this sequence version replaced gi:29469599.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

Center code: MIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L28535
 Center clone name: 326_C_3

Only the first 143.8 Kilobases of this clone are being submitted.
 The remainder overlaps accessions number AC138230 [WICR project
 L29013].

FEATURES

Source Location/Qualifiers
 1. 143835
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-326C3"
 /clone_id="RP11-11 Human Male BAC"
 1712..2006
 /rpt_family="AluSg"
 2347..2703
 /rpt_family="THB1B"
 4157..4302
 /rpt_family="L1MB4a"
 4435..4455
 /rpt_family="AT_rich"
 4528..4591
 /rpt_family="AT_rich"
 4592..4901
 /rpt_family="AluSx"
 4902..4968
 /rpt_family="AT_rich"
 4979..5091
 /rpt_family="L1MB4a"
 5154..5251
 /rpt_family="L1MB4a"
 5620..5683
 /rpt_family="L1MB4a"
 complement(5684..5954)
 /rpt_family="AluJo"
 5955..6050
 /rpt_family="L1MB4a"
 complement(6059..6206)
 /rpt_family="MER46C"
 6349..6578
 /rpt_family="L1MB4a"
 6792..7203
 /rpt_family="L2"
 complement(8884..9052)
 /rpt_family="L1MC5"
 9225..9287
 /rpt_family="MER68-int"
 9676..9986
 /rpt_family="AluSg"
 9994..10032
 /rpt_family="AT_rich"
 10182..10417
 /rpt_family="L1MB4"
 complement(10418..10946)
 /rpt_family="L1R59"
 complement(10947..11251)
 /rpt_family="AluSx"
 complement(11252..11272)
 /rpt_family="L1R59"
 complement(11290..11595)
 /rpt_family="AluSg"
 11702..12014
 /rpt_family="AluSg"
 12047..12137
 /rpt_family="L1M4"
 complement(12159..12465)
 repeat_region

/rpt_family="AluSx"
 12466..12605
 /rpt_family="L2"
 12763..12832
 /rpt_family="L2"
 12898..13715
 /rpt_family="L1MB5"
 13716..14007
 /rpt_family="AluSc"
 14018..14320
 /rpt_family="AluSx"
 14321..14603
 /rpt_family="L1MB5"
 14611..14772
 /rpt_family="AluSg/x"
 15386..15685
 /rpt_family="AluSx"
 15701..15984
 /rpt_family="AluJo"
 complement(16654..16949)
 /rpt_family="AluSx"
 17027..17095
 /rpt_family="L1MC5"
 complement(18948..19258)
 /rpt_family="AluY"
 19299..19409
 /rpt_family="L2"
 complement(19426..19719)
 /rpt_family="AluSg"
 complement(19742..20040)
 /rpt_family="AluJo"
 20207..20435
 /rpt_family="L1MC/D"
 20906..21032
 /rpt_family="(CCCCCG)n"
 complement(21181..21341)
 /rpt_family="MIR"
 complement(22663..22828)
 repeat_region

Query Match 68.6%; Score 446.6; DB 9; Length 143835;
 Best Local Similarity 99.1%; Pred.No.2.3e-78;
 Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	199	CTCCATCAAGCCCGAGATCAGAGGTGTGTGACTGTGAACGAGCCCGGCTTTGG	258
DB	83267	CTCTCTGACAGCCCGAGATCAGAGGTGTGTGACTGTGAACGAGCCCGGCTTTGG	83208
QY	259	CTCCAAAGCCAGTGTCTCAACATCTGTGCGCGATGTGACGCTGTGCGCCACTGCT	318
DB	83207	CTCCAAAGCCAGTGTCTCAACATCTGTGCGCGATGTGACGCTGTGCGCCACTGCT	83148
QY	319	GCTCCTGGGCTGTGTGTGACTGTGCTGCTGACCTGACCTGCGCCGCTGGCCAAAGACTGCG	378
DB	83147	GCTCCTGGGCTGTGTGTGACTGTGCTGCTGACCTGACCTGCGCCGCTGGCCAAAGACTGCG	83088
QY	379	CGCTCTTTCAGACCAAGTTTATGACGCGACTATGACTGACAGGCTGGTCTGATC	438
DB	83087	CGCTCTTTCAGACCAAGTTTATGACGCGACTATGACTGACAGGCTGGTCTGATC	83028
QY	439	TGGGGCACTAGCCCGACGACACTGACCCCAAGCTGCTGCTGGGGCCCAATATGACT	498
DB	83027	TGGGGCACTAGCCCGACGACACTGACCCCAAGCTGCTGCTGGGGCCCAATATGACT	82968
QY	499	CCCGAGAGCTGGGCGCTCTTGTGTGGGCTGCATCCCTGCGCCATCTGATCTGGGG	558
DB	82967	CCCGAGAGCTGGGCGCTCTTGTGTGGGCTGCATCCCTGCGCCATCTGATCTGGGG	82908
QY	559	CCCTCAGCCCCCAATGAGGCACTTAAGGCTGAACCAAGTCAACCCCGGGGCTTCAACC	618
DB	82807	CCCTCAGCCCCCAATGAGGCACTTAAGGCTGAACCAAGTCAACCCCGGGGCTTCAACC	82848
QY	619	TAAACGAGATTCCGGGGCCCTAATCTGCCC	651

DB 82847 TAACCCGAGAGTCCCGGCCCTAAGCTGCC 82815

RESULT 3
AP006283
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-326C3 map 11p, WORKING DRAFT
SEQUENCE, 19 unordered pieces.
ACCESSION AP006283
VERSION AP006283.1 GI:29243343
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA of 11p
JOURNAL Published Only in Database (2003)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: Humdract11
Center clone name: RP11-326C3
Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 144706 bases at least Q40
Consensus quality: 150967 bases at least Q20
Consensus quality: 153657 bases at least Q20
Insert size: 155019; sum-of-contigs
Quality coverage: 6.36x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1
21102 21001 contig of 21001 bp in length
40714 40714 contig of 19613 bp in length
58430 58430 contig of 17616 bp in length
58531 58531 contig of 8224 bp in length
66855 66855 contig of 14280 bp in length
81134 81134 contig of 8529 bp in length
89763 89763 contig of 12026 bp in length
101889 101889 contig of 10636 bp in length
101990 101990 contig of 7077 bp in length
112726 112726 contig of 5279 bp in length
119903 119903 contig of 3903 bp in length
125282 125282 contig of 7976 bp in length
129184 129184 contig of 7976 bp in length
137260 137260 contig of 6182 bp in length
143543 143543 contig of 3249 bp in length
146891 146891 contig of 3249 bp in length
149568 149568 contig of 2577 bp in length
152679 152679 contig of 3010 bp in length
152778 152778 gap of 100 bp
154136 154136 contig of 1357 bp in length
154236 154236 gap of 100 bp
154436 154436 gap of 100 bp
155662 155662 contig of 1427 bp in length
155763 155763 gap of 100 bp
156819 156819 contig of 1057 bp in length

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
21001: contig of 21001 bp in length
21002 21001: gap of 100 bp
21102 21101: contig of 19613 bp in length
40714 40714: gap of 100 bp
40715 40814: gap of 100 bp
58430 58430: contig of 17616 bp in length
58431 58530: gap of 100 bp
58531 58530: gap of 8224 bp in length
66755 66854: gap of 100 bp
66855 66854: gap of 100 bp
81135 81134: contig of 14280 bp in length
81135 81234: gap of 100 bp
89763 89763: contig of 8529 bp in length
89764 89863: gap of 100 bp
89864 89863: gap of 100 bp
101889 101889: contig of 12026 bp in length
101890 101889: gap of 100 bp
101990 101990: gap of 100 bp
112625 112625: contig of 10636 bp in length
112726 112725: gap of 100 bp
119802 119802: contig of 7077 bp in length
119903 119902: gap of 100 bp
125182 125181: contig of 5279 bp in length
125282 125281: gap of 100 bp
129185 129184: contig of 3903 bp in length
129185 129284: gap of 100 bp
129285 129284: gap of 100 bp
137261 137260: contig of 7976 bp in length
137261 137360: gap of 100 bp
137361 143542: contig of 6182 bp in length
143543 143642: gap of 100 bp
143643 146891: contig of 3249 bp in length
146892 146991: gap of 100 bp
146992 149568: contig of 2577 bp in length
149569 149668: gap of 100 bp
149669 152678: contig of 3010 bp in length
152679 152778: gap of 100 bp
152779 154135: contig of 1357 bp in length
154136 154235: gap of 100 bp
154236 155662: contig of 1427 bp in length
155663 155762: gap of 100 bp
155763 156819: contig of 1057 bp in length.

FEATURES
source
1..156819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p"
/clone="RP11-326C3"
1..21001
/note="assembly_fragment"
21102..40714
/note="assembly_fragment"
40815..58430
/note="assembly_fragment"
58531..66754
/note="assembly_fragment clone_end:SP6 vector_side:left"
66855..81134
/note="assembly_fragment"
81235..89763
/note="assembly_fragment"
89864..101889
/note="assembly_fragment"
101990..112625
/note="assembly_fragment"
112726..119802
/note="assembly_fragment"
119903..125181
/note="assembly_fragment"

```

misc_feature 125282..129184
               /note="assembly_fragment_clone_end:T7_vector_side:right"
misc_feature 129285..137260
               /note="assembly_fragment"
misc_feature 137361..143542
               /note="assembly_fragment"
misc_feature 143643..146891
               /note="assembly_fragment"
misc_feature 146992..149568
               /note="assembly_fragment"
misc_feature 149669..152678
               /note="assembly_fragment"
misc_feature 152779..154135
               /note="assembly_fragment"
misc_feature 154236..155662
               /note="assembly_fragment"
misc_feature 155763..156819
               /note="assembly_fragment"
BASE COUNT 37089 a 39987 c 40186 g 37757 t 1800 others
ORIGIN

```

```

Query Match      68.6%; Score 446.6; DB 2; Length 156819;
Best Local Similarity 99.1%; Pred. No. 2.3e-78;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Oy 199 CTCATCAAGGCCCGAGATCAGAAAGTGTGTGACCTGGAAGCGCCGCGCTTTTGG 258
Db 92819 CTCCTGAGCGCCCGAGATCAGAAAGTGTGTGACCTGGAAGCGCCGCGCTTTTGG 92878

Oy 259 CTCMAAGCCAGAGCTCAACATCTCTGCGCGGATGTGAGCGCTGTGCTCCGCACTGCT 318
Db 92879 CTCMAAGCCAGAGCTCAACATCTCTGCGCGGATGTGAGCGCTGTGCTCCGCACTGCT 92938

Oy 319 GCTCTGAGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 378
Db 92939 GCTCTGAGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 92998

Oy 379 GCGCTTCTTCAAGCCAGATTTGATGACGCGGATGATGATGATGATGATGATGATGATGATG 438
Db 92999 GCGCTTCTTCAAGCCAGATTTGATGACGCGGATGATGATGATGATGATGATGATGATGATG 93058

Oy 439 TGGGGCACTAAGCCCGAGACACTGACCCCGAGGCTGCTCCCTGCGCGGCAATCTGACT 498
Db 93059 TGGGGCACTAAGCCCGAGACACTGACCCCGAGGCTGCTCCCTGCGCGGCAATCTGACT 93118

Oy 499 CCGCGAGCGCTGCGCTCTTCTGTGAGGCGCTCCATCCCTGCGGATGATGATGATGATGATG 558
Db 93119 CCGCGAGCGCTGCGCTCTTCTGTGAGGCGCTCCATCCCTGCGGATGATGATGATGATGATG 93178

Oy 559 CCGTTCAGCCCGCAATGAGGCACTAAGCTGAACCAATGAGCCCGGCTCTTCAACC 618
Db 93179 CCGTTCAGCCCGCAATGAGGCACTAAGCTGAACCAATGAGCCCGGCTCTTCAACC 93238

Oy 619 TAACCGAGAGTTCCCGGCGCTTAATCTGCCC 651
Db 93239 TAACCGAGAGTTCCCGGCGCTTAATCTGCCC 93271

```

```

RESULT 4
AP006286/c 175416 bp DNA linear HTG 26-MAR-2003
LOCUS Homo sapiens chromosome 11 clone RP13-25N22 map 11p, WORKING DRAFT
DEFINITION
SEQUENCE, 22 unordered pieces.
ACCESSION AP006286
VERSION AP006286.1 GI:29243346
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

```

```

TITLE Homo sapiens genomic DNA of 11p
JOURNAL Published Only in Database (2003)
REFERENCE 2 (bases 1 to 175416)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: HumDra111
Center clone name: RP13-25N22

```

```

----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator RT-ampliphase; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 160074 bases at least Q40
Consensus quality: 167722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 173316; sum-of-contigs
Quality coverage: 7.92x in Q20 bases; sum-of-contigs

```

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 33131 contig of 33131 bp in length
33232 54098 contig of 20667 bp in length
54199 72037 contig of 17639 bp in length
72138 83277 contig of 11140 bp in length
83378 98599 contig of 15222 bp in length
98700 105513 contig of 16814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
122065 127627 contig of 5563 bp in length
127728 131971 contig of 4244 bp in length
132072 137506 contig of 5435 bp in length
137607 143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4509 bp in length
153249 157509 contig of 4261 bp in length
157610 160561 contig of 2952 bp in length
160662 164778 contig of 4117 bp in length
164879 168269 contig of 3391 bp in length
168370 170689 contig of 2320 bp in length
170790 172833 contig of 2044 bp in length
172934 174155 contig of 1222 bp in length
174256 175416 contig of 1161 bp in length.
NOTE: This is a 'working draft' sequence. It currently
consists of 22 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 33131: contig of 33131 bp in length
33132 33231: gap of 100 bp
33232 54098: contig of 20667 bp in length
54099 54198: gap of 100 bp
54199 72037: contig of 17639 bp in length
72038 72137: gap of 100 bp
72138 83277: contig of 11140 bp in length

```

* 83378 83377: gap of 100 bp
* 83378 98599: contig of 15222 bp in length
* 98600 98699: gap of 100 bp
* 98700 10513: contig of 6814 bp in length
* 10514 10513: gap of 100 bp
* 105614 116229: contig of 10616 bp in length
* 116230 116229: gap of 100 bp
* 116330 121964: contig of 5635 bp in length
* 121965 122064: gap of 100 bp
* 122065 127627: contig of 5563 bp in length
* 127628 131971: contig of 4244 bp in length
* 131972 132071: gap of 100 bp
* 132072 137506: contig of 5435 bp in length
* 137507 137607: gap of 100 bp
* 137607 143829: contig of 6223 bp in length
* 143830 143929: gap of 100 bp
* 143930 148539: contig of 4610 bp in length
* 148540 148639: gap of 100 bp
* 148640 153148: contig of 4509 bp in length
* 153149 153248: gap of 100 bp
* 153249 157509: contig of 4261 bp in length
* 157510 157609: gap of 100 bp
* 157610 160561: contig of 2952 bp in length
* 160562 160661: gap of 100 bp
* 160662 164778: contig of 4117 bp in length
* 164779 164878: gap of 100 bp
* 164879 168269: contig of 3391 bp in length
* 168270 170689: contig of 2320 bp in length
* 170690 170789: gap of 100 bp
* 170790 172833: contig of 2044 bp in length
* 172834 172933: gap of 100 bp
* 172934 174155: contig of 1222 bp in length
* 174156 174255: gap of 100 bp
* 174256 175416: contig of 1161 bp in length.

FEATURES

source
1. 175416
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p"
/clone="RP13-25N22"
1. 33131
/note="assembly_fragment clone_end:SP6 vector_side:left"
3332. 54098
/note="assembly_fragment"
54199. 72037
/note="assembly_fragment"
72138. 83277
/note="assembly_fragment"
83378. 98599
/note="assembly_fragment"
98700. 10513
/note="assembly_fragment clone_end:T7 vector_side:left"
105614. 116229
/note="assembly_fragment"
116330. 121964
/note="assembly_fragment"
122065. 127627
/note="assembly_fragment"
127728. 131971
/note="assembly_fragment"
132072. 137506
/note="assembly_fragment"
137607. 143829
/note="assembly_fragment"
143930. 148539
/note="assembly_fragment"
148640. 153148
/note="assembly_fragment"
153249. 157509

/note="assembly_fragment"
157610. 160561
/note="assembly_fragment"
160662. 164778
/note="assembly_fragment"
164879. 168269
/note="assembly_fragment"
168370. 170689
/note="assembly_fragment"
170790. 172833
/note="assembly_fragment"
172934. 174155
/note="assembly_fragment"
174256. 175416
/note="assembly_fragment"
BASE COUNT 37167 a 49392 c 49417 g 37339 t 2201 others
ORIGIN
Query Match 68.6%; Score 446.6; DB 2; Length 175416;
Best Local Similarity 99.1%; Pred. No. 2.2e-78;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 199 CTCATCAAGCCCGAGATCAGAAAGGTGCTGTAACCTGGAACCGGCGCTTTGG 258
Db CTCTCTGACAGCCCGAGATCAGAAAGGTGCTGTAACCTGGAACCGGCGCTTTGG 42150
QY 259 CTCGAAGCCCAATGCTACACATCTCTGCGCGGATGTGACCGTGGCCACTGCT 318
Db CTCGAAGCCCAATGCTACACATCTCTGCGCGGATGTGACCGTGGCCACTGCT 42090
QY 319 GCTCTGGGGCTGTGTGTGATCTGTGCTCTGACCTGGCCCGGCTGGCCAGGACTCTGC 378
Db GCTCTGGGGCTGTGTGTGATCTGTGCTCTGACCTGGCCCGGCTGGCCAGGACTCTGC 42030
QY 379 CGCCTCTTCAGACCAAGTTGATGATGACGGGACTATGACTGACAGCTGGGCTTCGATC 438
Db CGCCTCTTCAGACCAAGTTGATGATGACGGGACTATGACTGACAGCTGGGCTTCGATC 41970
QY 439 TGGGGCACTAGCCCGACAGACACTGACCCGAGCTGCTGCCCTGGGGCCCAATGACT 498
Db TGGGGCACTAGCCCGACAGACACTGACCCGAGCTGCTGCCCTGGGGCCCAATGACT 41910
QY 499 CCCCGAGCTGGGCTCTCTCTGTGTGGGCTTCATCTCCCTGATCTGTGGGG 558
Db CCCCGAGCTGGGCTCTCTCTGTGTGGGCTTCATCTCCCTGATCTGTGGGG 41850
QY 559 CCCTCAAGCCCAACAATGGGCACTTAAGCTGAACGACGACCCCGGGGTTCACCC 618
Db CCCTCAAGCCCAACAATGGGCACTTAAGCTGAACGACGACCCCGGGGTTCACCC 41790
QY 619 TAACCGAGAGTCCCGGCGCTTAATCTGCCC 651
Db TAACCGAGAGTCCCGGCGCTTAATCTGCCC 41757

RESULT 5
AC138230/c 187160 bp DNA linear PRI 22-MAY-2003
LOCUS Homo sapiens chromosome 11, clone RP13-317D12, complete sequence.
DEFINITION AC138230
ACCESSION AC138230.5 GI:30984765
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 187160)
AUTHORS Birren,B., Nusbaum,C. and Lander,B.
TITLE Homo sapiens chromosome 11, clone RP13-317D12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187160)
AUTHORS Birren,B., Nusbaum,C., Lander,B., Ali,A., Allen,N., Anderson,S.,
Barna,N., Basteien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,

Camata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Haez, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,
 Matthews, C., McCarthy, M., Meldrum, J., Menues, L., Mihova, T.,
 Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N.,
 Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K.,
 Travers, A., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome
 Research, 330 Charles Street, Cambridge, MA 02141, USA

TITLE
 JOURNAL
 DIRECT SUBMISSION

REFERENCE
 AUTHORS

Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome
 Research, 330 Charles Street, Cambridge, MA 02141, USA

TITLE
 JOURNAL
 DIRECT SUBMISSION

REFERENCE
 AUTHORS

Submitted (22-MAY-2003) Whitehead Institute/MIT Center for Genome
 Research, 330 Charles Street, Cambridge, MA 02141, USA

On May 22, 2003 this sequence version replaced gi:29423936.
 All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L29013
 Center clone name: 317_D_12

FEATURES
 source

Location/Qualifiers
 1..187160
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP13-317D12"
 /clone_11b="RP13-317D12 Human Female BAC"
 135..200
 /rpt_family="L2"
 1134..1310
 /rpt_family="WIR"
 complement(1311..1614)
 /rpt_family="AluY"
 1615..1638
 /rpt_family="WIR"
 1849..1938
 /rpt_family="L1ME"
 complement(14297..4385)
 /rpt_family="L2"
 4463..4561
 /rpt_family="FRPM"
 complement(4573..4613)
 /rpt_family="WDE1"
 4733..5040
 /rpt_family="AluY"
 complement(5094..5383)
 /rpt_family="L2"
 complement(5379..5505)
 /rpt_family="L2"
 complement(5501..5627)
 /rpt_family="L2"
 complement(5667..6096)
 /rpt_family="L2"
 6542..6666
 /rpt_family="MER41B"
 6699..6719
 /rpt_family="GC_rich"
 6917..6996
 /rpt_family="C-rich"
 7136..7329
 /rpt_family="C-rich"
 15461..15483
 /rpt_family="TGC)n"
 18178..18413
 /rpt_family="L2"
 complement(18543..18907)
 /rpt_family="MER4A"
 complement(18913..19138)
 /rpt_family="AluSg/x"
 complement(19139..19318)
 /rpt_family="MER4D"
 complement(19319..19688)
 /rpt_family="MER4B1"
 19689..19992
 /rpt_family="AluSp"
 19994..20308
 /rpt_family="AluSg"
 complement(20309..20630)
 /rpt_family="MER4B1"
 20632..20763
 /rpt_family="AluSg/x"
 complement(20785..20882)
 /rpt_family="MER4E"
 complement(20886..21091)

```

repeat_region      /rpt_family="MER41E"
21092..21402
/rpt_family="AlusX"
repeat_region      complement(21403..21580)
/rpt_family="MER41E"
repeat_region      complement(22261..22538)
/rpt_family="MER41E"
repeat_region      complement(22543..22594)
/rpt_family="MER41E"
repeat_region      22595..22902
/rpt_family="AluY"
repeat_region      complement(22903..23195)
/rpt_family="MER4D"
repeat_region      23197..23314
/rpt_family="FLAM_C"
repeat_region      complement(23318..23368)
/rpt_family="MER4D"
repeat_region      23369..23890
/rpt_family="MER4A"
repeat_region      complement(23891..24166)
/rpt_family="MER4D"
repeat_region      complement(24381..24698)
/rpt_family="MER63"
repeat_region      complement(25372..25668)
/rpt_family="AluY"
repeat_region      26402..26493
/rpt_family="(TG)n"
repeat_region      complement(27325..28044)
/rpt_family="LTR30"
28747..28755
/note="<30 qual SNGL region"
repeat_region      29024..29047
/rpt_family="GC_rich"
repeat_region      complement(29397..30103)
/rpt_family="LTR8"
30537..30656
/rpt_family="LTR2C"
repeat_region      complement(30657..31083)

Query Match      68.6%; Score 446.6; DB 9; Length 187160;
Best Local Similarity 99.1%; Pred. No. 2.2e-78;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 199 CTCACATCAGAGCCCGGATCAGAGGCTGTTGCTGACCTGGAACCGCGCGCTTTGG 258
Db 1606 CTCCTGCAAGCCCGGATCAGAGGCTGTTGCTGACCTGGAACCGCGCGCTTTGG 16547
QY 259 CTCCAAAGCCAAAGTGTACCAATCTGCGCCGATGTGAGCGCTGCGCCACTGCT 318
Db 16546 CTCCAAAGCCAAAGTGTACCAATCTGCGCCGATGTGAGCGCTGCGCCACTGCT 16487
QY 319 GCTCTGAGGCTGTGTGATGCTGTGCTGACCTGCGCCGCTGCGCCAAAGACTCTGC 378
Db 16486 GCTCTGAGGCTGTGTGATGCTGTGCTGACCTGCGCCGCTGCGCCAAAGACTCTGC 16427
QY 379 CGCTTCTTCAAGACCAAGTTTATGACCGGACCTTACTGACAGGCTGGTCTGATC 438
Db 16426 CGCTTCTTCAAGACCAAGTTTATGACCGGACCTTACTGACAGGCTGGTCTGATC 16367
QY 439 TGGGGCACTAGCCCGAGACACTAGACCCCAAGCTGCTGCGCCCGGAGCCCAATAGTACT 498
Db 16366 TGGGGCACTAGCCCGAGACACTAGACCCCAAGCTGCTGCGCCCGGAGCCCAATAGTACT 16307
QY 499 CCGCGAGCTGAGCTCTCTTCTGTGGGGCTTCATCCTGCGCCCAATCTGATCTGGGG 558
Db 16306 CCGCGAGCTGAGCTCTCTTCTGTGGGGCTTCATCCTGCGCCCAATCTGATCTGGGG 16247
QY 559 CCTTCAAGCCCAAGCTGAGCACTAAGCTGAAACAGATCAACCCCGGGAGTTTCAACC 618
Db 16246 CCTTCAAGCCCAAGCTGAGCACTAAGCTGAAACAGATCAACCCCGGGAGTTTCAACC 16187
QY 619 TAACCGAGAGTTCGGGCGCTTACTTGCC 651

```

```

Db 16186 TAACCGAGAGTTCGGGCGCTTACTTGCC 16154

RESULT 6
MMU9781      499 bp  mRNA  linear  ROD 22-MAR-2001
LOCUS      Mus musculus mRNA for putative haemopoietic membrane protein.
DEFINITION
VERSION      AJ009781.1 GI:3355654
KEYWORDS      haemopoietic membrane protein; putative.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS      Baird,J.W., Ryan,K.M., Hayes,I., Hampson,L., Heyworth,C.M.,
Clark,A., Wootton,M., Ansell,J.D., Menzel,U., Hole,N. and
Graham,G.J.
TITLE      Differentiating embryonal stem cells are a rich source of
JOURNAL      haemopoietic gene products and suggest erythroid preconditioning of
MEDLINE      primitive haemopoietic stem cells
PUBMED      J. Biol. Chem. 276 (12), 9189-9198 (2001)
21153738
PUBMED      11106657
2 (bases 1 to 499)
AUTHORS      Graham,G.J.
TITLE      Direct Subnasion
JOURNAL      Submitted (24-JUL-1998) Graham G.J., CRC Beatson Laboratories,
Beatson Institute for Cancer Research, Gartcube Estate, Switchback
Road, Bearsden, Glasgow G61 1BD, SCOTLAND, UK

FEATURES
source
1..499
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/codon_start=1
/product="putative haemopoietic membrane protein"
/protein_id="CA08829.1"
/db_xref="GI:3355655"
/cb_xref="SPTREMBL:O88728"
/translation="MDTSYPRDPAPSSRKDAADAAHTALSMGTPPTPRDHLWGVF
STMYNLCCIGFALVHSVKARDQKAGLLEAARQYGSAAKCNILAAWMTLVPIPLL
LDLVYTGALHLKAKDAFAFSTKPFDEBDYN"

BASE COUNT      107 a 155 c 137 g 100 t

CDS
1..436
/note="ORF1"
/codon_start=1
/product="putative haemopoietic membrane protein"
/protein_id="CA08829.1"
/db_xref="GI:3355655"
/cb_xref="SPTREMBL:O88728"
/translation="MDTSYPRDPAPSSRKDAADAAHTALSMGTPPTPRDHLWGVF
STMYNLCCIGFALVHSVKARDQKAGLLEAARQYGSAAKCNILAAWMTLVPIPLL
LDLVYTGALHLKAKDAFAFSTKPFDEBDYN"

Query Match      43.7%; Score 284.6; DB 10; Length 499;
Best Local Similarity 79.6%; Pred. No. 4.8e-46;
Matches 351; Conservative 0; Mismatches 84; Indels 6; Gaps 1;

QY 1 GGAAGAGAGCGGCTGGAACCCATGAGACAGGCGTATCCCGGAGAGACCCGGGCCCC 60
Db 9 GGAAGAGAGCGGCTGGAACCCATGAGACAGGCGTATCCCGGAGAGACCCGGGCCCC 68
QY 61 CACGCCCAAGAGCGCGT-----GCCACACAGCCCTCAACATGCGCGCCCCGACCC 114
Db 69 ATATCTCCGCAAGGCTGATGCTGAGCCACACAGCCCTCTCAATGAGAAACCTGGCCC 128
QY 115 CCGGCTCGAGACCACTTATGCTGCTGTGTTAGACACCTCTAATCTGATGTTG 174
Db 129 TACACCAAGAGTACATGCTGTGCTGTCTTCAAGCAAGATGTAATCTGTGCTG 188
QY 175 CCTGGCTTCTGCGGCTGCTCACTCAATCAAGCCCGAGATCAAGAGTGGTGTGA 234
Db 189 CTTGGAATTCCTGCGGCTGCTCACTGTCTCAAGCCCGAGACCAAGAGATGGCTGGAA 248
QY 235 CTGGAAGAGCGCGGCTTTGGCTTCAAGCCCAAGGCTCAATCTGCGCGCAT 294

```

Db 249 CTGGAGGCTGCAAGGAGTATGCTCCAAAGCCAACTGCTACACATCTCTGGCTGCAT 308
 QY 295 GTGAGCAGCTGTGTCGGCCACTGCTCTCTGCGGGCTGTGATGATGCTGTCACCT 354
 Db 309 GTGAGCAGCTGTGTCGGCCACTGCTCTCTGCGGGCTGTGATGATGCTGTCACCT 368
 QY 355 GGGCCGCTGTGTCGGCCACTGCTCTCTGCGGGCTGTGATGATGCTGTCACCT 414
 Db 369 GTGAGCAGCTGTGTCGGCCACTGCTCTCTGCGGGCTGTGATGATGATGAGGACTA 428
 QY 415 TGAATGACAGCTGTGTCCTG 435
 Db 429 TAACTAAGATTCGAGCTG 449

RESULT 7
 AC132968/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-172C3, *** SEQUENCING IN PROGRESS

AC132968 238533 bp DNA linear HTG 20-NOV-2002
 AC132968 4 GI:25139219
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Buxarota; Metaxoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 238533)
 Wuzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bialwalo, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gubaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoque, M.,
 Hollins, B., Howells, S., Hulys, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensbawa, L., Loulsegod, H., Lozada, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwachukwu, O., Okunnu, G., Olamunsgoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plappier, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,
 Piazzi, B., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojia, A., Rose, R., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Slisom, I., Sitter, C. D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanai, K.,
 Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,

Williams, G., Wilson, R., Wlarczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., Von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 238533)
 Reference
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (05-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 238533)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

Genome Center

Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: XCMS
 Center clone name: CH230-172C3

Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 22317 bases at least Q40
 Consensus quality: 22343 bases at least Q30
 Consensus quality: 22655 bases at least Q20
 Estimated insert size: 230654; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently
 consists of 4 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 222461: contig of 222461 bp in length
 * 222462 222561: gap of unknown length
 * 222562 224949: contig of 2386 bp in length
 * 224950 225049: gap of unknown length
 * 225050 228242: contig of 3193 bp in length
 * 228243 228342: gap of unknown length
 * 228343 238533: contig of 10191 bp in length.

FEATURES

source

1.238533
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-172C3"
 1. 2349
 /note="wgs_contig"

misc_feature

```

misc_feature      124155..209379
                  /note="clone_boundary
                  clone_end:T7
                  site:
                  end_sequence:BH290527"
BASE COUNT      61451 a 53521 c 52085 g 60005 t 11471 others
ORIGIN
Query Match      26.6%; Score 173; DB 2; Length 238533;
Best Local Similarity 81.6%; Fred. No. 1.4e-24;
Matches 200; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 201 CCATCAAGGCCGAGATCAGAAAGGTTGGAGCTTGGAGCGGCCGCGGCTTTGGCT 260
DB 213892 CCTCGAGGCCCGAGACAGAGAGTGAAGCCGGGACCTTGAGGCTGCAAGAGATGAGCT 213833
QY 261 CCAAGCCAAAGTGTATCAACATCTCTGCGCGGATGTGAGACCTGTGCGCCACTGTGTC 320
DB 213832 CCAAGCCAAAGTGTATCAACATCTCTGCGCGGATGTGAGACCTGTGCGCCACTGTGTC 213773
QY 321 TCCCTGGGCTGTGTGTGATCTGTGCTCCCTGCACTTGCGCCGCTGCGCCAGAGACTGTGCG 380
DB 213772 TCCCTGGGCTGTGTGTGATCTGTGCTCCCTGCACTTGCGCCGCTGCGCCAGAGACTGTGCG 213713
QY 381 CCTCTTCAGGACCAAGTTTATGATGACGCGGACTATGACTGACAGGCTGGGTCTTGATCTG 440
DB 213712 CTTCTTCAGGACCAAGTTTATGATGAGAGACTATGACTGAGATTCTTGACCTGTGCTG 213653
QY 441 GAGGA 445
DB 213652 AACCA 213648

RESULT 8
AC109272/c      200734 bp DNA linear HTG 23-APR-2003
LOCUS           Mus musculus clone RP23-354L18, WORKING DRAFT SEQUENCE, 17 ordered
DEFINITION      pieces.
ACCESSION      AC109272.3 GI:30181119
VERSION         HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS        Mus musculus (house mouse)
SOURCE          Mus musculus
ORGANISM        Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1 (bases 1 to 200734)
                Birren,B., Nusbaum,C. and Lander,E.
                Mus musculus, clone RP23-354L18
                Unpublished
                2 (bases 1 to 200734)
                Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
                Anderson,S., Barta,N., Baetsen,V., Boguslavsky,L., Boukhalter,B.,
                Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
                Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
                Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
                Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
                Grinde,S., Gord,S., Goyette,W., Graham,L., Grand-Pierre,N.,
                Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
                Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
                Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
                MacDonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M.,
                McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T., Norbu,C.,
                Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
                Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
                Peterson,K., Phunhkhang,P., Pierre,N., Pollara,V., Raymond,C.,
                Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
                Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
                Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
                Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
                Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
                Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
                Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
                Direct Submission
TITLE

```

```

JOURNAL
REFERENCE
AUTHORS
Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 200734)
Birren,B., Nusbaum,C., Lander,E., Aboueleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barta,N., Baetsen,V., Bloom,T.,
Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Grande,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N.,
Rachpka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2003 this sequence version replaced g1:20330961.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RN/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L20944
Center clone name: 354_L_18

NOTE: This is a 'working draft' sequence. It currently
consists of 17 contigs. Gaps between the contigs
are represented as runs of 'N'. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.
This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
1 5430: contig of 5430 bp in length
5431 5530: gap of 100 bp
5531 6612: contig of 1082 bp in length
6613 6712: gap of 100 bp
6713 8249: contig of 1537 bp in length
8250 8349: gap of 100 bp
8350 10942: contig of 2593 bp in length
10943 11042: gap of 100 bp
11043 14087: contig of 3045 bp in length
14088 14187: gap of 100 bp
14188 20506: contig of 6319 bp in length
20507 20607: gap of 100 bp
20608 26335: contig of 5729 bp in length
26336 26435: gap of 100 bp
26436 31423: contig of 4988 bp in length
31424 31523: gap of 100 bp
31524 39064: contig of 7541 bp in length
39065 39164: gap of 100 bp
39165 47891: contig of 8766 bp in length
47892 47991: gap of 100 bp
47992 56258: contig of 8268 bp in length
56259 56358: gap of 100 bp
56359 66903: contig of 10545 bp in length
66904 67004: gap of 100 bp
109182: contig of 42179 bp in length

```

```

* 109183 109282: gap of 100 bp
* 109283 133485: contig of 24203 bp in length
* 133486 133586: gap of 100 bp
* 133586 157263: contig of 23678 bp in length
* 157264 157363: gap of 100 bp
* 157364 192898: contig of 35535 bp in length
* 192899 200734: gap of 100 bp
* 192999 200734: contig of 7736 bp in length.
Location/Qualifiers
1. 200734
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-354L18"
/clone_1b="RPC1-23 Female Mouse BAC"
1. 5430
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
5531..6612
/note="assembly_fragment"
6713..8249
/note="assembly_fragment"
8350..10942
/note="assembly_fragment"
11043..14087
/note="assembly_fragment"
14188..20506
/note="assembly_fragment"
20607..26335
/note="assembly_fragment"
26436..31423
/note="assembly_fragment"
31524..33964
/note="assembly_fragment"
33965..47890
/note="assembly_fragment"
47891..56258
/note="assembly_fragment"
56359..66903
/note="assembly_fragment"
67004..109182
/note="assembly_fragment"
109283..133485
/note="assembly_fragment"
133586..157263
/note="assembly_fragment"
157364..192898
/note="assembly_fragment"
192899..200734
/note="assembly_fragment"
clone_end:T7
vector_side:right"
misc_feature
50239 a 48869 c 48384 g 51632 t 1610 others
BASE COUNT
ORIGIN
Query Match 25.7%; Score 167.6; DB 2; Length 200734;
Best Local Similarity 63.8%; Pred. No. 1.6e-23;
Matches 277; Conservative 0; Mismatches 144; Indels 13; Gaps 1;

```

```

Db 144930 CTTCTTCAGACCAAGTTGATGAGAGAGACTATTAAGATTCCGAGCTG----- 144876
Qy 441 GGGACATAGCCCGACGACACTGACCCCGAGCTGCTGCTGCGGCGCAATACGACTCC 500
Db 144875 -----TCCTGAACCGAGGACAAACATGTAGTACGCTTGCCCAACACGACTCC 144824
Qy 501 CCGAGCTGAGCCCTCTCTCTGTGAGGAGCTTCACCTCCCGCCATCTGAGGCC 560
Db 144823 TGGAGATTACAGCTACTATAGACACCTTCCTCCCGAGCGGTAGCTCAGAGGGCGC 144764
Qy 561 CTCAGCGCCCAATAGGACCTTAAGCTGTAACCACTCAGACCCGGGCTTACACCTTA 620
Db 144763 CCTCTGTCCACGATGTGATTTGATGCTGCGCCCACTAGATCTTCAGACTTAATTAT 144704
Qy 621 ACCCGAGATTGCC 634
Db 144703 AACCTAGAGGCTCC 144690

RESULT 9
AC107815/c
LOCUS
DEFINITION
Mus musculus clone RP23-114A6, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
AC107815
VERSION
AC107815.4 GI:31442494
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 251206)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-114A6
Unpublished
2 (bases 1 to 251206)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barina,N., Bastien,V., Boguslavsky,L., Bouhgaltier,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chararo,B.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Lander,T., Lehoczký,J., Levine,R., Liu,G., Maclean,C.,
Macedonia,P., Major,C., Margis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneses,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Punthang,P., Pierre,N., Pollara,V., Raymond,C.,
Reiter,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupback,R., Seaman,S.,
Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Sudramanian,A., Talams,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,D., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submissions
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 251206)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barina,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Bouhgaltier,B., Camarata,J., Chang,J., Chapel,Y.,
Collymore,A., Cook,A., Cooke,P., Cornu,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hateg,N., Hateg,N., Hateg,N.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,

```


TITLE
JOURNAL
COMMENT

Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkham, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rice, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Teftaye, S., Theodore, V., Topnam, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Submitted (06-JUN-2003). Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 6, 2003 this sequence version replaced gi:20163115.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Center project name: L20545

Center clone name: 114_A6

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; Version 0.960731

Consensus quality: 250046 bases at least Q40

Consensus quality: 250590 bases at least Q20

Consensus quality: 250806 bases at least Q20

Insert size: 230000; agarose-1p

Insert size: 250906; sum-of-coverage

Quality coverage: 11.6 in Q20 bases; agarose-1p

Quality coverage: 11.6 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 29459: contig of 29459 bp in length
* 29460 29559: gap of 100 bp
* 29560 30528: contig of 969 bp in length
* 30529 30628: gap of 100 bp
* 30629 142773: contig of 112145 bp in length
* 142774 142874: gap of 100 bp
* 142874 251206: contig of 108333 bp in length.

FEATURES
source

1. 251206
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_11b="RPC1-23 Female Mouse BAC"
1. 29459
/note="assembly_fragment"
clone_end:SP6
vector_side:left
29560..30528
/note="assembly_fragment"
30629..142773
/note="assembly_fragment"
142874..251206
/note="assembly_fragment"
clone_end:77
vector_side:right

BASE COUNT 67395 a 56774 c 57538 g 69199 t 300 others

ORIGIN

Query Match 25.7% Score 167.6; DB 2; Length 251206;
Best Local Similarity 63.8%; Pred. No. 1.6e-23;
Matches 277; Conservative 0; Mismatches 144; Indels 13; Gaps 1;

201 CCATCAAGGCCCGGAGATGAGAGAGGTGTGTGACCTTGAAGCCGCGCTTTTGCT 260
131688 CCTCAGAGCCCGGAGACAGAGAGATGCTGGGAACCTTGAAGGCTCAAGGAGTATGGCT 131629
261 CCAAGCCCAAGTGTCTCAACAACATCTCGGCGCGGATGTGACGCTGTGCCCACTGTCC 320
131628 CCAAGCCCAAGTGTCTCAACAACATCTCGGCGCGGATGTGACGCTGTGCCCACTGTCC 131569
321 TCTTGGGCGTGTGTGTGACGCTGTGCGCTGTGACCTGCGCGGCTGTGCGCAAGACTGTGCG 380
131568 TCTTGGGCGTGTGTGTGACGCTGTGCGCTGTGACCTGCGCGGCTGTGCGCAAGACTGTGCG 131509
381 CTTTCTTCAAGCAAGTTTATGACGCGGACCTATGACTGACAGGCTGTGATCTG 440
131508 CTTTCTTCAAGCAAGTTTATGACGCGGACCTATGACTGACAGGCTGTGATCTG 131454
441 GGGCACTAGCCCGGAGACACGACACCCGAGGCTGTGCGCGGCGGCAATATGACTGCC 500
131453 -----TCCCTGAACCGAGGACACACCATGTCAGTGTGCTCCCAACACGACTCC 131402
501 CCGAGCGTGTGCGCTCTTGTGTGCGGCGCTGCATCCCTGCCCATCTGATCTGTGCGGCGC 560
131401 TGGAGGTATGACGCTTACTATGACACCTGATCCCTGCCCATCTGATCTGTGCGGCGC 131342
561 CTCAGCCCGGCAATGGGACCTTAAAGGCTGAACGATGACACCCGCGGCTTTCACCTTA 620
131341 CCTCTTGTCAAGATGTATTTATGATGCTGCCACATGATCCCTTCACTTAATTAT 131282
621 ACCGAGAGTTC 634
131281 AACCTAAGGCTTC 131268

RESULT 10
AC006970 136098 bp DNA linear PRI 27-SEP-2000

LOCUS Homo sapiens clone RP4-725G10, complete sequence.
DEFINITION AC006970
ACCESSION AC006970.6 GI:10312289
KEYWORDS HTG.

SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 136098)
Waterson, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
2 (bases 1 to 136098)
Waterson, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE
AUTHORS 3 (bases 1 to 136098)
Waterson, R.H.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT
On Sep 27, 2000 this sequence version replaced gi:9838025.
Center project name: H.DU0725G10.
Location/Qualifiers
1. 136098
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

Web site: <http://www.sbgc.stanford.edu>
 Contact: (Jackson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Kodiriquez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAC Plate: 14 Row: a Column: 11
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10835237.
 Location/Qualifiers

FEATURES

source
 1..678
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /db_xref="locusID:10581"
 /clone="MGC:9196 IMAGE:3876542"
 /tissue_type="Lung, large cell carcinoma"
 /clone_id="NIH MGC_68"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 31..429
 /codon_start=1
 /product="interferon induced transmembrane protein 2 (1-8D)"
 /protein_id="AAH09696.1"
 /db_xref="GI:16307215"
 /translation="NMHIVTFSFSPNSGPPNYEMKEQEVAMLAGPHNPAPSTV
 IHRSETPVDHVMISLFTNFTTCCLGFIARAYSVKSRDRKMGVDVTGAQAYASTA
 KCLINWALILGIMTILVLIPLVAVOAG"

CDS

BASE COUNT 146 a 231 c 148 g 153 t
 ORIGIN

Query Match 14.3%; Score 93.4; DB 9; Length 678;
 Best Local Similarity 62.8%; Pred. No. 1.8e-08;
 Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATCTGCGGTGCTCAGACACCTTACTGATGTTGCTCCGCTTC 184
 DB 193 GACCACTGCTGCGGTGCTCAGACACCTTACTGATGTTGCTCCGCTTC 252
 QY 185 CTGGCGCTGCTCTCTCATCAAGCCGAGATCAGAAAGTGTGTCACCTGAAGG 244
 DB 253 ATAGCACTGGCCCTACTCCGTAAGTCTAGAGACAGAAAGTGTGTCACCTGAAGG 312
 QY 245 GCCCGCGCTTTGGCTCCAAAGCCAAAGTCTACACATCTCTGCGCGAGTGTGACGCTG 304
 DB 313 GCCCAGGCGCTATGCTCCACCGCCAAAGTCTGAAACATCTGGGCGCTGATTTGGGCATC 372
 QY 305 GTGCCCGCACTGCTCTCTCTGCGGTGCTGACGTGCTCCCTGCACTG 355
 DB 373 TTCATGACCATCTGCTCTCTCATCCAGTGTGTGCTCCAGGCCCG 423

RESULT 13
 AX337050 905 bp DNA linear PAT 09-JAN-2002
 LOCUS
 DEFINITION Sequence 7559 from Patent WO0194629.
 ACCESSION AX337050
 VERSION AX337050.1 GI:18127769
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 Young, P. B., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D. R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 JOURNAL gene sets
 PATENT: WO 0194629-A 7559 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES Location/Qualifiers

source

1..905
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 189 a 304 c 206 g 206 t
 ORIGIN

Query Match 14.3%; Score 93.4; DB 6; Length 905;
 Best Local Similarity 62.8%; Pred. No. 1.7e-08;
 Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATCTGCGGTGCTCAGACACCTTACTGATGTTGCTCCGCTTC 184
 DB 442 GACCACTGCTGCGGTGCTCAGACACCTTACTGATGTTGCTCCGCTTC 501
 QY 185 CTGGCGCTGCTCTCTCATCAAGCCGAGATCAGAAAGTGTGTCACCTGAAGG 244
 DB 502 ATAGCACTGGCCCTACTCCGTAAGTCTAGAGACAGAAAGTGTGTCACCTGAAGG 561
 QY 245 GCCCGCGCTTTGGCTCCAAAGCCAAAGTCTACACATCTCTGCGCGAGTGTGACGCTG 304
 DB 562 GCCCAGGCGCTATGCTCCACCGCCAAAGTCTGAAACATCTGGGCGCTGATTTGGGCATC 621
 QY 305 GTGCCCGCACTGCTCTCTCTGCGGTGCTGACGTGCTCCCTGCACTG 355
 DB 622 TTCATGACCATCTGCTCTCTCATCCAGTGTGTGCTCCAGGCCAG 672

RESULT 14

LOCUS HS18D 905 bp mRNA linear PRI 26-MAY-1993
 DEFINITION Human 1-8D gene from interferon-inducible gene family.
 ACCESSION X57351
 VERSION X57351.1 GI:311373
 KEYWORDS 1-8 gene family; 1-8D gene; interferon inducible gene.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 905)
 TITLE Lewis, A. R., Reid, L. E., McMahon, M., Stark, G. R. and Kerr, I. M.
 JOURNAL Molecular analysis of a human interferon-inducible gene family
 MEDLINE Eur. J. Biochem. 199 (2), 417-423 (1991)
 PUBMED 91301153
 REFERENCE 2 (bases 1 to 905)
 AUTHORS Kerr, I. M.
 TITLE Direct Submission
 JOURNAL Submitted (22-JAN-1991) I. M. Kerr, Imperial Cancer Research Fund,
 44 Lincoln's Inn Fields, London WC2A 3PX, U K
 On Jul 27, 1993 this sequence version replaced gi:23395.
 COMMENT See X02490 for overlapping cDNA sequence.
 The Human 1-8D gene shows sequence identity to 1-8U (see X57352)
 and 9-27 (see J04164 for cDNA sequence, X02491 for cDNA sequence).

FEATURES

source
 1..905
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="lymphoid"
 /clone_id="loristB"
 1..905
 /gene="1-8D"
 /gene="1-8D"
 189..202
 /note="ISRS (interferon-stimulable response element)"
 /evidence=experimental
 280..678
 /gene="1-8D"
 /codon_start=1
 /protein_id="CAA0625.1"
 /db_xref="GI:23396"
 /db_xref="SWISS-PROT:Q01629"

```

exon
/translation="MNHIVQTFSPVNSGPPNYEMLKEEDEVANLGGPHNPAPPTSTV
IHIRETSVDPHVWVSLFNTLFPNMCICGFIAPFAYSVKSRDKWGDVTGAQAVASTA
KCLNIWALIGIFMTLIVILVILVIVLVOAQR"
<280..626
/number=1
/number=1
/evidence=experimental
627..2678
/number=2
/number=2
/evidence=experimental
698..765
/number=1
/number=1
/evidence=experimental
766..833
/number=1
/number=1
/evidence=experimental
886..891
/number=1
/number=1
/evidence=experimental
repeat_unit
/number=1
/number=1
/evidence=experimental
repeat_unit
/number=1
/number=1
/evidence=experimental
polya_signal
/number=1
/number=1
/evidence=experimental
BASE COUNT 189 a 304 c 206 g 206 t
ORIGIN
Query Match 14.3%; Score 93.4; DB 9; Length 905;
Best Local Similarity 62.8%; Pred. No. 1.7e-08;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
125 GACCACTGATCTGCGGTGTTGACGACCCCTCACTGATCTGCTGCTGCTTC 184
442 GACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
185 CTGCGCGTGGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 244
502 ATAGCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
245 GCCCGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304
562 GCCCAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
305 GTGCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 355
622 TTCTATGACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
RESULT 15
BT006892 402 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens interferon induced transmembrane protein 3 (1-8U)
DEFINITION mRNA, complete cds.
ACCESSION BT006892
VERSION BT006892.1 GI:30582622
KEYWORDS FLI CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 402)
AUTHORS Kalinine,N., Chen,X., Rolfe,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Pheasant,M., and Farmer,A.
TITLE Cloning of human full-length cDNAs in BD Creator (TM) System Donor
vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 402)
AUTHORS Kalinine,N., Chen,X., Rolfe,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Pheasant,M., and Farmer,A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA
COMMENT This cDNA clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the

```

```

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00117X1.0"
/clone_lib="BD Creator (TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/notes="Vector: pDNR-Dual"
1..402
/codon_start=1
/product="interferon induced transmembrane protein 3
(1-8U)"
/protein_id="AAP35538.1"
/db_xref="GI:30582623"
/translation="MNHIVQTFSPVNSGPPNYEMLKEEDEVANLGGPHNPAPPTST
VHIRETSVDPHVWVSLFNTLFPNMCICGFIAPFAYSVKSRDKWGDVTGAQAVAST
AKCLNIWALIGIFMTLIVILVILVIVLVOAQR"
BASE COUNT 77 a 140 c 98 g 87 t
ORIGIN
Query Match 14.3%; Score 93.2; DB 9; Length 402;
Best Local Similarity 66.3%; Pred. No. 2.2e-08;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
125 GACCACTGATCTGCGGTGTTGACGACCCCTCACTGATCTGCTGCTGCTTC 184
166 GACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225
185 CTGCGCGTGGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 244
226 ATAGCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
245 GCCCGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304
286 GCCCAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345
305 GTGCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 326
346 CTCATGACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 367

```

Search completed: January 5, 2004, 17:31:03
Job time : 2823 secs


```

PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 17479
LENGTH: 326
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006970.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 16
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 48
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 88
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 53
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1e+02
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 88
OTHER INFORMATION: EST HUMAN HIT: AA527771.1, EVALUATE 0.00e+00
OTHER INFORMATION: NT HIT: g111995467, EVALUATE 1.00e-129
OTHER INFORMATION: SWISSPROT HIT: Q01629, EVALUATE 3.00e-35
US-09-864-761-17479

Query Match 14.7%; Score 95.4; DB 9; Length 326;
Best Local Similarity 67.2%; Pred. No. 1.5e-15;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 126 ACCACTTGATCTGATGCGTGTTCAGACACCTCTACCTGATCTGTTGCTCGCTTCC 185
Db |||||
QY 209 ACCATGTTGTCGTGTCCTCCCTGTTCAACACCTCTTANTGAACCCCTGCTGGCTTCA 150
Db |||||
QY 186 TGGCGCTGCGCTTACTCTCATCAAGCGCCGAGATCAAGAAGTGTGTGATCTGGAACGG 245
Db |||||
QY 149 TAGGTTCACTCACTCATCAAGCTTAAGGACAGAAAGATGTGTGGCGACTGAACGGGG 90
Db |||||
QY 246 CCGGCGGTTTGTGCTCCAAAGCCAGTGTCAACATCTCGCCCGGAGTGAAGCTCG 305
Db |||||
QY 89 CCGAGGCTTATGCTCTCACCGCCAGAGTCTGGAACATCTGGGCTCTGACTTTGGGCATCC 30
Db |||||
QY 306 TGGCGCCACTGCTGCTCTCG 326
Db |||||
QY 29 TCATGACCATTTGCTCATCG 9
Db |||||

RESULT 2
US-09-864-761-692/c
Sequence 692, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Weishang
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456

```

```

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263, 6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 692
LENGTH: 459
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006970.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 16
OTHER INFORMATION: EXPRESSED IN HELLIO, SIGNAL = 48
OTHER INFORMATION: EXPRESSED IN PETAL LAYER, SIGNAL = 88
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 53
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1e+02
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 88
US-09-864-761-692

Query Match 14.7%; Score 95.4; DB 9; Length 459;
Best Local Similarity 67.2%; Pred. No. 1.5e-15;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

126 ACCACTTGATGTGGTGGATGTTGACGACACCTCTACCTGATCTGTGTTGCTCGGCTTCC 185
Db ACCATGTTGTCCTGCTCCCTGTTCAACACCCCTCTCATGAAACCCCTGCTGCTGGGGCTTCA 347
QY TGGGGCTGGGCTACTCATCATCAAGCCGGAAGTCAAGAGTGTTGGTGGTGAAGCGG 245
Db TAGGCTTCACTTACCTCATGAGCTTAGGAGACAGAAAGATGTTGGGCACTTGACCGGGG 287
QY CCGGCGCTTTGGCTCCAAAGCCAAAGTCTAACACATCTTGCGCGCAGATGTGACGCTGG 305
Db CCGAGGCGCTATGCTCCACCGCCAAAGTGGCTGAACATCTGGGCGCTTGAGCATCC 227
QY TGGCGGCACTGCTGCTCTGG 326
Db TCAATGACCATTCGTGCTCATG 206

```

```
RESULT 3
US-10-197-666A-67
; Sequence 67, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABI SYR
; TITLE OF INVENTION: E1k1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(389)
US-10-197-666A-67

Query Match
Best Local Similarity 14.3%; Score 93.4; DB 15; Length 621;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATGTCGTGGTGTTCAGACACCTTACTGATGTCGTGGCTTC 184
DB 156 GACCACTGATGTCGTGGTGTTCAGACACCTTACTGATGTCGTGGCTTC 215
QY 185 CTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244
DB 216 ATAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 275
QY 245 GCCCGCGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304
DB 276 GCCCGCGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 335
QY 305 GTGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 355
DB 336 TTCATGACCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386

RESULT 4
US-09-969-708-88
; Sequence 88, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
```

```
SOFTWARE: PatentIn version 3.0
; SEQ ID NO 88
; LENGTH: 905
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-88

Query Match
Best Local Similarity 14.3%; Score 93.4; DB 10; Length 905;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATGTCGTGGTGTTCAGACACCTTACTGATGTCGTGGCTTC 184
DB 442 GACCACTGATGTCGTGGTGTTCAGACACCTTACTGATGTCGTGGCTTC 501
QY 185 CTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244
DB 502 ATAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
QY 245 GCCCGCGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304
DB 562 GCCCGCGCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
QY 305 GTGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 355
DB 622 TTCATGACCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672

RESULT 5
US-10-301-822-82
; Sequence 82, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kametkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen J.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEM01-0292RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 905
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (280)..(678)
US-10-301-822-82

Query Match
Best Local Similarity 14.3%; Score 93.4; DB 13; Length 905;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATGTCGTGGTGTTCAGACACCTTACTGATGTCGTGGCTTC 184
DB 442 GACCACTGATGTCGTGGTGTTCAGACACCTTACTGATGTCGTGGCTTC 501
QY 185 CTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244
```


TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (280) ..(675)
US-10-197-666A-145

Query Match 14.3%; Score 93.4; DB 15; Length 905;
Best Local Similarity 62.8%; Pred. No. 5.1e-15;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATCTGTCGGTGTTCAGACCCCTTACCTGAATCTGTTGCTCCGCTTC 184
DB 442 GACCACTGTCGTGTCCTGTCCTGTTCAACCCCTTTATGAACCTGCTGCTGCTTC 501
QY 185 CTGGCGCTGGCTTACTTCATCAAGCCCGAGATCAAGAGTGTGTGACCTGGAAGC 244
DB 502 ATAGCATTCGCTACTCCGTAAGTCTAAGGACAGAAAGTGTGGCCAGCGTGAACCGG 561
QY 245 GCCCGCGTTTGGCTTCCCAAGCAAGTGTACAAATCTGGCCGCGATGTGACGCTG 304
DB 562 GCCCAGGCTATGCTTCCACCGCAAGTGTGACATCTGGGCTGTGATTTGGGCAATC 621
QY 305 GTGCCGCACTGCTGCTCCGCTGAGTGTGACTGTGCTGACCTGACCTG 355
DB 622 TTCATGACCATTCCTGCTGCTCATCATCCAGTGTGTGCTGCAAGCCGAG 672

RESULT 9

US-10-197-666A-147
Sequence 147, Application US/1019766A
Publication No. US20030092037A1
GENERAL INFORMATION:

APPLICANT: ASAMI KASEI KABUSIKI KAISYA
TITLE OF INVENTION: E1K1 phosphorylation related gene
FILE REFERENCE: PH-1548US
CURRENT APPLICATION NUMBER: US/10/197,666A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: JP 2001-218204
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: JP 2001-263450
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: JP 2002-012176
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: US 60/305,884
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/316,304
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 60/350,027
PRIOR FILING DATE: 2002-01-23
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 147
LENGTH: 905
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (280) ..(675)
US-10-197-666A-147

Query Match 14.3%; Score 93.4; DB 15; Length 905;
Best Local Similarity 62.8%; Pred. No. 5.1e-15;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATCTGTCGGTGTTCAGACCCCTTACCTGAATCTGTTGCTCCGCTTC 184
DB 442 GACCACTGTCGTGTCCTGTCCTGTTCAACCCCTTTATGAACCTGCTGCTGCTTC 501
QY 185 CTGGCGCTGGCTTACTTCATCAAGCCCGAGATCAAGAGTGTGTGACCTGGAAGC 244
DB 502 ATAGCATTCGCTACTCCGTAAGTCTAAGGACAGAAAGTGTGGCCAGCGTGAACCGG 561

QY 245 GCCCGCGTTTGGCTTCCAAAGCAAGTGTACACATCTTGGCCGCGATGTGACGCTG 304
DB 562 GCCCAGGCTATGCTTCCACCGCAAGTGTGACATCTGGGCTGTGATTTGGGCAATC 621
QY 305 GTGCCGCACTGCTGCTCCGCTGAGTGTGACTGTGCTGACCTGACCTG 355
DB 622 TTCATGACCATTCCTGCTGCTCATCATCCAGTGTGTGCTGCAAGCCGAG 672

RESULT 10

US-10-115-831-27
Sequence 27, Application US/10115831
Publication No. US20030219743A1
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyang
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
FILE REFERENCE: Polypeptides
FILE REFERENCE: 792CIP2ADIV
CURRENT APPLICATION NUMBER: US/10/115,831
CURRENT FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 09/667,298
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 178
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 27
LENGTH: 920
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (280) ..(678)
US-10-115-831-27

Query Match 14.3%; Score 93.4; DB 13; Length 920;
Best Local Similarity 62.8%; Pred. No. 5.1e-15;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATCTGTCGGTGTTCAGACCCCTTACCTGAATCTGTTGCTCCGCTTC 184
DB 442 GACCACTGTCGTGTCCTGTCCTGTTCAACCCCTTTATGAACCTGCTGCTGCTTC 501
QY 185 CTGGCGCTGGCTTACTTCATCAAGCCCGAGATCAAGAGTGTGTGACCTGGAAGC 244
DB 502 ATAGCATTCGCTACTCCGTAAGTCTAAGGACAGAAAGTGTGGCCAGCGTGAACCGG 561
QY 245 GCCCGCGTTTGGCTTCCAAAGCAAGTGTACACATCTTGGCCGCGATGTGACGCTG 304
DB 562 GCCCAGGCTATGCTTCCACCGCAAGTGTGACATCTGGGCTGTGATTTGGGCAATC 621
QY 305 GTGCCGCACTGCTGCTCCGCTGAGTGTGACTGTGCTGACCTGACCTG 355
DB 622 TTCATGACCATTCCTGCTGCTCATCATCCAGTGTGTGCTGCAAGCCGAG 672

RESULT 11

US-10-198-846-10182
Sequence 10182, Application US/10198846
Publication No. US20030099974A1
GENERAL INFORMATION:

APPLICANT: Li, Y. James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinhilber, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049

PRIOR APPLICATION NUMBER: US 60/305,884
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/316,304
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 60/350,027
PRIOR FILING DATE: 2002-01-23
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 71
LENGTH: 606
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (45)..(443)
US-10-197-666A-71

Query Match 14.3%; Score 93.2; DB 15; Length 606;
Best Local Similarity 66.3%; Pred. No. 5.6e-15;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 125 GACCACTTGATCTGTGGTTCAGACCCCTTACCTGATCTGTGCTGCTTC 184
DB 210 GACCACTTGATCTGTGGTTCAGACCCCTTACCTGATCTGTGCTGCTTC 269
QY 185 CTGGCGCTGGCTTACTTCATCAAGCCGAGATCAGAGGTGTTGTGACTGGAAGCG 244
DB 270 ATAGCATTCGCTTACCTCCGAGATCTGAGGACAGAGATGTTGGCGACGTGACCGGG 329
QY 245 GCCCGCGTTTGGCTCCAAAGCCAGTGTACCAATCCGGCCGAGATGTGGAAGCG 304
DB 330 GCCCAGGCTTATGCTTCCACCGCCCAATGCTGAAATCTGGGCTTGAATTCGGGCATC 389
QY 305 GTGCCGCGCACTGCTGCTCTCG 326
DB 390 CTCATGACCATCTGCTCATCG 411

RESULT 15
US-10-197-666A-75
Sequence 75, Application US/10197666A
Publication No. US20030092037A1
GENERAL INFORMATION:
APPLICANT: ASAHI KASEI KABUSIKI KAISYA
TITLE OF INVENTION: E1k1 phosphorylation related gene
FILE REFERENCE: PH-1548US
CURRENT APPLICATION NUMBER: US/10/197,666A
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: JP 2001-218204
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: JP 2001-263450
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: JP 2002-012176
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: US 60/305,884
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/316,304
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 60/350,027
PRIOR FILING DATE: 2002-01-23
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 75
LENGTH: 621
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (60)..(458)
US-10-197-666A-75

Query Match 14.3%; Score 93.2; DB 15; Length 621;
Best Local Similarity 66.3%; Pred. No. 5.6e-15;

Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 125 GACCACTTGATCTGTGGTTCAGACCCCTTACCTGATCTGTGCTGCTTC 184
DB 225 GACCACTTGATCTGTGGTTCAGACCCCTTACCTGATCTGTGCTGCTTC 284
QY 185 CTGGCGCTGGCTTACTTCATCAAGCCGAGATCAGAGGTGTTGTGACTGGAAGCG 244
DB 285 ATAGCATTCGCTTACCTCCGAGATCTGAGGACAGAGATGTTGGCGACGTGACCGGG 344
QY 245 GCCCGCGTTTGGCTCCAAAGCCAGTGTACCAATCCGGCCGAGATGTGGAAGCG 304
DB 345 GCCCAGGCTTATGCTTCCACCGCCCAATGCTGAAATCTGGGCTTGAATTCGGGCATC 404
QY 305 GTGCCGCGCACTGCTGCTCTCG 326
DB 405 CTCATGACCATCTGCTCATCG 426

Search completed: January 5, 2004, 19:03:06
Job time : 298 secs